Phase Two Report from the Working Group on Student Learning Outcomes and Assessment

March 2011

Phase Two Report from the
Working Group on Student Learning Outcomes and Assessment

Contents

Working Group Membership

- **Dr. Lois Alves,** Vice President for Enrollment Services, Research, and Planning, Middlesex Community College

- **Dr. Mario Borunda,** Dean of Education, Lesley University, Member, Board of Higher Education

- **Dr. Patricia Crosson, Chair;** Professor Emeritus of Higher Education, University of Massachusetts Amherst; Senior Advisor for Academic Policy, Department of Higher Education

- **Dr. Michael Fiorentino,** Executive Vice President / Academic Affairs, Fitchburg State University

- **Dr. Lane Glenn,** Vice President of Academic Affairs, Northern Essex Community College

- **Dr. Kate Harrington,** Associate Vice President, Academic Affairs, Student Affairs, and International Relations, University of Massachusetts President's Office

- **Liam Hogan,** Parliamentarian of the Student Senate, Holyoke Community College

- **Dr. Jonathan Keller,** Associate Commissioner for Research, Planning, and Information Systems, Massachusetts Department of Higher Education

- **Dr. Nicholas Massa,** Professor of Engineering Technology, Springfield Technical Community College

- **Dr. Patricia Maguire Meservey,** President, Salem State University

- **Dr. Christopher J. O'Donnell,** Professor of Mathematics, Massachusetts Maritime Academy

- **Dr. Patricia Plummer,** Senior Advisor to the Senior Vice President for Academic Affairs, Student Affairs, and International Relations, University of Massachusetts President's Office

- **Dr. Robert Pura,** President, Greenfield Community College

- **Dr. Michael Young,** Associate Provost for Academic Planning and Administration, Bridgewater State University

Former members: Both served until June 2010

- *Lois Martin, Professor of Mathematics, Retired, Massasoit Community College*

- *Julianne Cormio, Student Board Representative, Framingham State University*

Executive Summary

The Working Group on Student Learning Outcomes and Assessment (WGSLOA) was established by Richard Freeland, Commissioner of Higher Education, in late fall 2009 in anticipation of the Vision Project, a bold effort to set a public agenda for higher education and commit public campuses and the Department/Board of Higher Education to producing nationally leading educational results at a time when the need for well-educated citizens and a well-prepared workforce is critical for the future of the Commonwealth. Commissioner Freeland asked the Working Group to concentrate on the student learning aspect of what would become the Vision Project and especially on the assessment of student learning. Learning outcomes assessment has become a national movement that aims to more clearly understand and demonstrate what students know and are able to do as a result of their college experiences. The Working Group brings together presidents, provosts, faculty members, institutional research or assessment directors, Department of Higher Education staff, a member of the Board of Higher Education, and students from public colleges and universities in Massachusetts for an in-depth and collaborative examination of student learning outcomes and assessment.

WGSLOA submitted its first report to Commissioner Freeland in July 2010. It focused on learning outcomes and assessment programs on the public college and university campuses in Massachusetts. (Appendix A contains the Executive Summary of the Phase One report.)

This Phase Two report draws upon findings and recommendations from the first report and reflects a diligent effort on the part of the Working Group to respond to the Phase Two charge. Commissioner Freeland asked the WGSLOA to consider a number of state-level issues related to learning outcomes assessment—accountability, comparability, transparency, tests, and rankings; to provide an analysis or where Massachusetts stands in comparison to other states with respect to policies and practices; and to provide a plan for system-wide learning outcomes assessment based on a set of common learning outcomes and assessment measures that builds on and supports strong campus-based assessment programs.

The Working Group has done its best to fulfill the Commissioner's charge. Ten meetings were held during Phase Two, and Working Group members spent many additional hours reading and corresponding. The group also attended conferences, met with leading experts in the field and hosted two faculty conversations on learning outcomes assessment. The WGSLOA was very impressed by the level of faculty sophistication, engagement and interest in assessment, and its members benefited a great deal from the nearly 200 faculty members who attended. Many faculty observations are contained in this report.

The Working Group believes that our report and Plan for Statewide Learning Outcomes Assessment reflects best thinking and best practice in learning outcomes assessment and responds to the needs of the Vision Project. Our plan represents the collective effort and responsibility of the community colleges, state universities, the University of Massachusetts and the Department of Higher Education. It speaks to the collective desire of public higher education to improve student learning and learning outcomes assessment and to the willingness to be held accountable for doing so. It recognizes accountability to executive and legislative leaders and the taxpayers who provide essential resources but also to students, parents, alumni and other stakeholders, and ultimately, to ourselves for the quality of the work that we do.

The first chapter of this report begins with discussion of the importance of learning outcomes assessment and its relationship to accountability in higher education. Although tensions between assessment and accountability have long been recognized in higher education, more recent work emphasizes the importance of, and possibilities for, reconciling these tensions. The chapter also treats the long-standing practice of quality assurance provided through regional accrediting bodies, including the New England Association for Schools and Colleges (NEASC) and its Commission on Institutions of Higher Education (CIHE). The Working Group believes that NEASC standards and peer-review-based processes and the new CIHE emphasis in the area of student assessment serve Massachusetts public institutions well and deserve our support. We also recognize and support the Vision Project agenda for public higher education in Massachusetts with its commitment to quality and to measure, compare and report publicly on student learning. We believe that the plan we propose in this report will allow colleges and universities to respond to both.

The report also describes the difficulties associated with developing appropriate institutional groupings and metrics for comparability from campus to campus, segment to segment, state to state, and with national benchmarks. Comparisons can be helpful for improvement and accountability for campuses and for the system, but we want to stress the importance of "doing it right." The WGSLOA believes that peer institutional and segment-based peer-state comparisons of student learning outcomes would provide the most useful information. We encourage a formal review of peer institutions for each campus, the inclusion of student learning as a category for comparison and the careful selection of appropriate learning metrics. Partnering with "Massachusetts-like" states to compare student learning outcomes would provide better information than is now available from efforts to compare the 50 states. We suggest that a dashboard, accompanied by customized reports to contextualize the information in different ways to respond to the interests and needs of various audiences, would be a good way of presenting assessment results.

The WGSLOA recommends the following four principles as a guide to policy and practice in Massachusetts. The WGSLOA believes these guiding principles will help to manage the tensions that exist between assessment for improvement and assessment for accountability, preserve and strengthen institutional capacity for evidence-based continuous improvement, and provide the accountability important for public support of public higher education. The principles are taken from the work of Peter Ewell (2009):

1. ***Respond visibly to domains of legitimate external concern.***
2. ***Show action on the results of assessment.***
3. ***Emphasize assessment at the major transition points in a college career.***
4. ***Embed assessment in the regular curriculum.***

Chapter Two of the report recognizes the challenges inherent in the seemingly contradictory imperatives for improvement and accountability, privacy and transparency, and uniqueness and comparability and argues that an effective system-wide plan for learning outcomes assessment must nonetheless allow room for all of these dimensions and bring them into appropriate balance. It points out that research and literature on sound assessment practice is clear that no single instrument or approach to assessing learning can meet all of the challenges and notes that the most effective systemic models of assessment offer multiple measures of student learning in a triangulation approach. This reinforces a position taken in the WGSLOA's Phase One report that

there should be no high-stakes standardized assessment instrument for all students in the Commonwealth's higher education system.

Chapter Two also introduces a conceptual model and presents a Plan for Statewide Learning Outcomes Assessment. In the plan, "LEAP Essential Learning Outcomes" provide a flexible framework for robust campus-based assessment because the LEAP learning outcomes speak directly to the qualities that we want our undergraduate students to possess. At the campus level, the plan includes curriculum-based embedded assessments, direct assessments involving such things as professional licensure or graduate/professional school admissions exams or standardized tests when selected by programs or campuses, and indirect assessments such as the Community College Survey of Student Engagement (at two-year colleges) or the National Survey of Student Engagement (at four-year college and university campuses).

Campus assessment work enables system-level features such as a multiple measures report presentation or dashboard for the Vision Project. Additionally, learning measures could be combined into a composite learning indicator. State partnerships would provide opportunities to compare learning outcomes results among similar states.

The final chapter of the report speaks to issues of resources and implementation. It recognizes that new or reallocated resources will be needed for campuses and the system to improve the capacity for evidence-based assessment through continued refinement of learning outcomes, curricular alignment activities, use of multiple and embedded assessment measures, greater attention to using results for teaching and learning improvement, and more widespread assessment reporting.

The Plan for Statewide Learning Outcomes Assessment relies heavily on extensive faculty involvement and a professional assessment staff. Resources in support of both will be needed. The Working Group believes that faculty and staff involvement (including staff in areas of academic and student affairs, assessment and institutional research) is essential to achieving lasting, positive impact on student learning. While assessment is a component of the work of faculty and designated professionals, building a statewide model cannot be accomplished by placing additional expectations on personnel whose current work responsibilities are already extensive. For each segment (i.e., community college, state university, the University of Massachusetts) and for the system as a whole, overall faculty size, faculty teaching loads and collective bargaining contracts will need to be made consistent with expectations in the area of assessment. Other activities that will incur a cost are those associated with professional development for faculty and staff and technology costs associated with web-based assessment platforms. New networks for cross-campus collaboration and sharing of best practices should help to improve practices and reduce overall costs although this too will require faculty and staff time and effort.

Successful implementation for the statewide plan is dependent upon the active participation of all undergraduate public higher education institutions in the Commonwealth and on their willingness to join with DHE in collaborative efforts. The plan also requires DHE to recognize the work that is already underway at the campus level and to allow for the uniqueness of each campus in the system to be valued in assessment work. The final chapter recommends that the Commissioner, with the assistance of the advisory body of presidents, take several implementation steps including:

- Seek to become a LEAP State and establish a Massachusetts LEAP Team, with representation from each segment (the University of Massachusetts, state universities, and community colleges) and DHE to work on LEAP State issues and initiatives.

- Develop a Resource Plan that outlines direct and indirect resources needed by DHE and campuses to achieve the goals of the Vision Project and an implementation timeline that responds to priorities and available resources

- Engage campus provosts or chief academic officers to provide leadership for faculty engagement and campus assessment activity; LEAP-related analysis, application and endorsement; and other activities in support of the statewide model for learning outcomes assessment.

- Establish a network of faculty and professional staff assessment leaders representing each campus and DHE to engage in a series of collaborative activities designed to strengthen campus programs and implement the statewide plan. This work will occasionally involve institutional research colleagues as well.

Ultimately, improved capacity and commitment to assessment on campuses, the strength of collaborative work across the public campuses, active participation from DHE and strong support for the Vision Project by public higher education stakeholders will ensure the success of this important effort. The final paragraph in Chapter One also serves well as an overall conclusion for this summary:

> *The Vision Project provides the opportunity to create a system of assessment that has intention to develop a culture of inquiry and improvement within the public higher education system; support existing efforts at each college and university; support faculty and staff in the development of their own learning about assessment; increase the system's accountability to future students, their families and the taxpayer; increase the accountability at each college and university to their own faculty, staff and students; and increase the confidence and commitment of key stakeholders and policy makers in the state public higher education system (Phase Two Report, p.18).*

Introduction

The Working Group on Student Learning Outcomes and Assessment (WGSLOA) was established by Richard Freeland, Commissioner of Higher Education, in late fall 2009 in anticipation of the Vision Project, a bold effort to set a public agenda for higher education and commit public campuses and the Department/Board of Higher Education to producing nationally leading educational results at a time when the need for well-educated citizens and a well-prepared workforce is critical for the future of the Commonwealth. Commissioner Freeland asked the Working Group to concentrate on the student learning aspect of what would become the Vision Project and especially on the assessment of student learning. Learning outcomes assessment has become a national movement that aims to more clearly understand and demonstrate what students know and are able to do as a result of their college experiences. The Working Group brings together presidents, provosts, faculty members, institutional research or assessment directors, Department of Higher Education staff, a member of the Board of Higher Education, and students from public colleges and universities in Massachusetts for an in-depth and collaborative examination of student learning outcomes and assessment.

WGSLOA submitted its first report to Commissioner Freeland in July 2010. It focused on learning outcomes and assessment programs on the public college and university campuses in Massachusetts (Appendix A). This Phase Two report draws upon findings and recommendations from the first report and reflects a diligent effort on the part of the Working Group to respond to our Phase Two charge. Commissioner Freeland asked us to focus more directly on state-level issues related to learning outcomes assessment and to:

- **Consider development of a set of student learning outcomes** that reflect the common aspirations of our campuses and consider the advantages of having all public campuses use them. Campuses, of course, would continue to have additional learning outcomes based on their own unique missions, programs and student bodies.

- **Review studies, reports and recommendations** by national and state higher education organizations that focus on state-level, as opposed to institutional, policies and best practices for the assessment of college student learning, including materials about state-mandated testing and state practices that link student learning assessment with statewide planning, performance measurement, public accountability, state ranking systems and international comparisons in higher education.

- Consistent with the outcome of the Board of Higher Education review of the Vision Project, **consider the most effective way to implement the outcome** that calls for Massachusetts to be a national leader in academic achievements of public college students on campus-level and national assessments of learning.

- **Consider issues of accountability and transparency** as they relate to student learning outcomes and assessment. Colleges, multi-campus systems and state boards vary in the content and amount of information from student learning assessment that is made public. What levels of accountability and transparency are appropriate for the Commonwealth?

The Commissioner asked the Working Group to provide a report at the end of Phase Two that:

- **Provides an analysis** of where Massachusetts stands in comparison to others in state efforts to support learning outcomes and assessment.
- **Speaks to issues** of state policy and practice, assessment tests, state rankings, accountability and transparency.
- **Provides a plan** for system-wide learning assessment based on a set of common learning outcomes and assessment measures that builds on and supports strong campus-based assessment programs.

We have done our best to fulfill our charge. During Phase Two we met ten times as a group and spent many additional hours reading and corresponding. We also conducted a session at the Vision Project Academic Conference, attended the annual forum of the New England Educational Assessment Network (NEEAN), and, joined by Commissioner Freeland, met with Dr. Peter Ewell, the Vice President of the National Center for Higher Education Management Systems (NCHEMS), a leading authority on learning outcomes assessment. Several members of the Working Group attended a LEAP States Summit meeting sponsored by the Association of American Colleges and Universities (AAC&U). This allowed us to delve more deeply into LEAP Learning Outcomes and Value Rubrics as a possible framework for a system-wide approach to learning outcomes assessment and to learn more about what LEAP State status would mean for the Commonwealth.

WGSLOA invited faculty members from public campuses and universities to join us for two informal conversations about learning outcomes assessment. We were impressed by the level of faculty sophistication, engagement and interest in assessment, and we benefited a great deal from the nearly 200 faculty members who attended. We reviewed all comments from these conversations and many are reflected in our observations and recommendations. We have also included a representative selection of faculty comments in shaded "views from the field" and "in the words of the faculty" sections throughout this report.

Our Phase One Report prompted the Department of Higher Education (DHE) to prepare a grant proposal for the Davis Educational Foundation. The Foundation's generous support will enable implementation of many of our recommendations including the hiring of a professional assessment expert to assist campuses in the review of their programs, the formation of a team of assessment leaders from all campuses, professional development for faculty and professional staff, and pilot work with assessment measures. This will provide an excellent beginning for the collaborative work suggested in this report.

This Phase Two report is presented in three chapters:

- **Chapter One** focuses on assessment of student learning from a state-level perspective and takes up the closely related issues of accountability, quality assurance and accreditation, comparability, transparency and rankings. It also examines assessment patterns and practices in other states.
- **Chapter Two** responds to the Commissioner's request for a plan for system-wide learning assessment based on a set of common learning outcomes and assessment measures and building from strong campus-based programs. We recommend a Plan for Statewide Learning Outcomes Assessment that responds to many challenges and builds from a conceptual model of how campuses and system can work together. The chapter includes

discussion of six plan components: LEAP as a common framework, Embedded Assessment, Direct Assessment, Indirect Assessment, a Composite Learning Indicator and State Partnerships.

- **Chapter Three** takes up issues of resources and implementation. It focuses attention on campus and system responsibility for learning outcomes assessment, mechanisms for collaboration across public campuses, connections to the Vision Project, and the many things that will be needed to turn the plan into reality. It recognizes the high ambitions of the Vision Project and of our Plan for Statewide Learning Outcomes Assessment and suggests a way forward that the Working Group believes is appropriate for Massachusetts and likely to lead to the leadership position for Massachusetts aspired to in the Vision Project.

Appendices to the Report include an executive summary of the Phase One Report of the WGSLOA, a list of LEAP Essential Learning Outcomes, definitions and a list of LEAP Value Rubrics, and a description of a Composite Indicator of Educational Capital. The composite indicator of educational capital was prepared by the WGSLOA as a potential state-level measure for use with the Vision Project but it is not included in the Plan for Statewide Learning Outcomes Assessment.

Chapter One
Learning Outcomes Assessment: From a State Perspective

In this chapter we focus on several multi-faceted issues that relate to how states interact with their public colleges and universities, other organizations, and other states in learning outcomes assessment. We focus first on Assessment and Accountability, providing definitions and describing the two paradigms of assessment for improvement and assessment for accountability. We discuss the tensions that result from them and endorse suggestions of Peter Ewell for moving beyond them. Ultimately, assessment and accountability are about quality, so we next take up the subject of quality assurance in higher education and the system of regional accreditation. We address issues of comparability, transparency and rankings, all of which are intended to respond to state-level concerns for accountability and are therefore important to state-level policy considerations for learning outcomes assessment. We review patterns and practices in other states and conclude with a recommendation for moving forward.

Assessment

> *"...institutional assessment efforts should not be concerned about valuing what*
> *can be measured but, instead, about measuring that which is valued."*
> (Astin, 1991, as cited by Banta, 1996, p. 5)

Assessment is the systematic collection of information about student learning, using the time, knowledge, expertise, and resources available, in order to inform decisions that affect student learning (Walvoord, 2010). Trudy Banta et al provide further clarity with regard to formative and summative assessment; "Summative assessment is demonstrating that our students have achieved certain levels, while formative assessment is used for improving student learning, teaching, and programs" (Banta, Griffin, Flateby & Kahn, 2009).

The learning assessment movement in higher education (as it is thought of today as something separate from faculty evaluation of student work for grading purposes) began in the early 1980s. Its etiology is rooted in curriculum reform reports that called for "greater curricular coherency, the use of powerful pedagogies known to be associated with high learning gains, and knowledge about student outcomes and experiences" (Ewell, 2009, p. 5). Some say it began even earlier, in the 1920s and 30s when state governments began using newly available tests and surveys to demonstrate to the business community and taxpayers that the investment in higher education was a wise one (Shavelson, 2010).

Today, the assessment of student learning has become a very important activity in higher education. Colleges and universities in Massachusetts and elsewhere have become extensively engaged in it, as documented in our Phase One report. Federal and state leaders, statewide governing and coordinating boards, business executives, accrediting bodies, parents and students, and the general public all want more and better information about what students have learned in college and what they can do with what they have learned. Different patterns and practices for state-level policy, coordination, and support of learning outcomes assessment have evolved over the last decades. WGSLOA had much to do to respond to the Commissioner's request that our Phase Two work focus on state-level issues in learning outcomes assessment.

WGSLOA believes that assessment is a fiber in higher education's cultural development, one in which the honest understanding of who we are leads to the dynamic discovery of who we can become in service to students, faculty, staff and community. For that to occur, assessment must be viewed as both a systematic as well as the systemic organization of tools and practices for inquiry and improvement.

Accountability

> *"There is inadequate transparency and accountability for measuring institutional performance, which is more and more necessary to maintaining public trust in higher education."*
> (U.S. Department of Education, *A Test of Leadership,* 2006, p. 13)

In "A Test of Leadership," what is now known as the Spellings Commission (after then-Secretary of Education Margaret Spellings) ushered in a new era of accountability. The influential federal report recognized the importance of higher education to the economy and well-being of the country and to each state; it also recognized the cost and as a result demanded to see the evidence of performance, including in student learning. It connected the assessment of student learning, perhaps irrevocably, to accountability including state-level accountability. It captured the growing sentiment of taxpayers, those elected officials who represent them, and the business community—in essence, the key stakeholders of public higher education.

> *"Our complex, decentralized postsecondary education system has no comprehensive strategy, particularly for undergraduate programs, to provide either adequate internal accountability systems or effective public information. Too many decisions about higher education—from those made by policymakers to those made by students and families—rely heavily on reputation and rankings derived to a large extent from inputs such as financial resources rather than outcomes. Better data about real performance and lifelong working and learning ability is absolutely essential if we are to meet national needs and improve institutional performance."* (U.S. Department of Education, 2006, p. 13)

> *"Despite increased attention to student learning results by colleges and universities and accreditation agencies, parents and students have no solid evidence, comparable across institutions, of how much students learn in colleges or whether they learn more at one college than another. Similarly, policymakers need more comprehensive data to help them decide whether the national investment in higher education is paying off and how taxpayer dollars could be used more effectively."* (U.S. Department of Education, 2006, p. 13)

In 2005, shortly before the Spellings Commission report was released, a National Commission on Accountability in Higher Education formed by the State Higher Education Executive Officers (SHEEO) released its own report on accountability. It noted:

> *"Real (and needed) improvement in higher education will come when accountability in higher education is a democratic process through which shared goals are explicitly established, progress is measured, and work to improve performance is motivated and guided. It will include: agreement on fundamental priorities, an effective, practical division of labor, focus on a few critical goals at every level of responsibility, and rigorous*

> *measurement and public reporting of results, followed by collaborative work to improve"* (National Commission on Accountability in Higher Education, 2005, p. 7).

The SHEEO National Commission described the problem of accountability not as an absence or deliberate omission on the part of higher education, since many accountability initiatives had been launched after Sputnik in 1957 and proliferated after a Nation at Risk in 1983, but as a "failure to implement accountability approaches that help improve performance in a complex, decentralized system of higher education" (p. 11). The Commission emphasized the need for fresh, new approaches fashioned by educators and policy makers. The Spellings and SHEEO reports, and other initiatives by state governors' associations, higher education associations, and state system leaders influenced state policy and institutional responses to the press for greater accountability for student learning. They brought us to the assessment movement of today.

The Working Group recognizes the connection between assessment of student learning and accountability. Assessment provides the necessary data and evidence for the advancement of student learning. Assessment improves learning. The assessment process also provides increased accountability to ourselves, our students, the public, and investors in higher education. In the final analysis, the WGSLOA agrees with the following statement: "Accountability requires that we take responsibility for the story we commit ourselves to telling" (Shulman, 2007).

Assessment and Accountability: Tensions and Paradigms

But it is the tensions and not the connections between assessment for improvement and assessment for accountability that capture the lion's share of attention in higher education literature and public debate.

> *"Too often accountability is a battleground between educators and policymakers. Many educators believe externally imposed accountability is a tool to place blame or avoid responsibility for inadequate financial support. Many policymakers, frustrated because existing investments are not producing better results, believe stronger external accountability is the only way to get improvement. In an atmosphere of resentment and mistrust, accountability initiatives produce more resistance than progress."* (National Commission on Accountability in Higher Education, 2005, p. 11)

The tensions result from different underlying perceptions about the purposes of assessment and different choices about how best to do it. Peter Ewell describes two distinct paradigms:

Table 1. Two Assessment Paradigms

	Assessment for Improving Paradigm	Assessment for Accountability Paradigm
Strategic Dimensions		
Intent	Formative (Improvement)	Summative (Judgment)
Stance	Internal	External

Predominant Ethos	Engagement	Compliance

Application Choices

Instrumentation	Multiple/Triangulation	Standardized
Nature of Evidence	Quantitative and Qualitative	Quantitative
Reference Points	Over Time, Comparative, Established Goal	Comparative or Fixed Standard
Communication of Results	Multiple Internal Channels and Media	Public Communication
Uses of Results	Multiple Feedback Loops	Reporting

Source: Ewell (2009). *Assessment, Accountability, and Improvement: Revisiting the Tension*, p. 8.

Assessment for Improvement; Assessment as Accountability: Guiding Principles

The seeds for the current tensions between assessment for improvement and assessment for accountability were planted over 30 years ago, and the tensions continue to be important in current debates. But much has changed over the years. Peter Ewell identifies the most important changes: greater perceived legitimacy for assessment by academics who realize that assessment has become "a condition of doing business in higher education"; recognition of the new policy centrality of higher education among governmental and business leaders; increasing emphasis on learning outcomes assessment by regional and specialized accrediting bodies and federal oversight bodies; and newly developed technology tools for assessment.

He concludes: *"Because the stakes associated with higher education are so much higher for policy makers today, aggressive action on the accountability agenda is more likely and a proactive response on the part of the academy is more urgent"* (Ewell, 2009, p. 7).

The WGSLOA recommends the following four principles as a guide to policy and practice in response to the calls for increased external accountability. The WGSLOA believes these guiding principles will help to manage the tensions that exist, help preserve and develop institutional capacity for evidence-based continuous improvement, and provide the accountability important for public support of public higher education. The principles are taken from the work of Peter Ewell (2009), and we have quoted or summarized his arguments here:

1. *Respond visibly to domains of legitimate external concern.* "… the current external demand is more about providing direct evidence of student academic achievement-- centered on broad undergraduate skills like critical thinking, communication, and problem solving than about where graduates go (p. 14)." Higher education's posture of engagement about assessment's accountability dimension today must emphasize a commitment to *a collective responsibility* for teaching and learning and their results. Instead of seeing assessment as an aspect of higher education's responsibility to its *funders*—legitimate though this may be—both faculty and academic leaders need to see

it as part of our accountability to *ourselves.* This is, after all, how we operate in the realm of research, and it is why mechanisms like independent peer review are so important to maintaining scholarly integrity. It needs to happen in teaching and learning as well (p. 15)."

2. ***Show action on the results of assessment.*** *"*Accordingly, colleges and universities will not only have to demonstrate sincere efforts to improve student learning but will also have to prove that their students are achieving at adequate levels in the first place. This will increasingly mean reporting actual learning outcomes in comparative or benchmarked forms as well as being transparent about internal efforts at continuous improvement… (p. 15-16)."

3. ***Emphasize assessment at the major transition points in a college career.*** *"*If an institution's goals for student learning are truly dominant, they must permeate the entire curriculum and be explicitly assessed at multiple points in a student's career. This means assessment for college readiness on entry and requiring clear demonstrations of mastery to earn a degree. It also requires attention to what happens between those two points... (p. 17)."

4. ***Embed assessment in the regular curriculum.*** Assessment needs to be systemically and systematically connected to teaching and learning processes and for too long they have been seen as a separate activity—for accreditation or state. "Embedded assessment approaches like competency-anchored assignments and electronic portfolios are becoming increasingly comfortable for faculty and can now be implemented efficiently on a large scale. They represent what is today the most compelling route that an institution can take to navigate the accountability-improvement tension (p. 20)."

Quality Assurance, Accreditation and the New England Association of Schools and Colleges (NEASC)

Assessment and Accountability share a fundamental interest in the quality of higher education. Public assurances about the quality of colleges and universities in the United States have long been provided through processes of "regional accreditation." Seven accreditation organizations operate in six geographic regions of the country through nongovernmental, non-profit voluntary associations and accredit nearly 3,000 colleges and universities. Accreditation is a self-regulatory, peer review process based on rigorous standards. Colleges and universities are judged based on self-evaluations analyzing how well they meet these standards, in light of their mission. Following a review by a team of peers, accrediting commissions determine the accreditation status of the institution and use a variety of means to ensure follow-up as appropriate and further evaluation in the case of substantive change on the part of the institution. Their approval provides a stamp of quality analogous to a "Good Housekeeping Seal," although it is far more serious in its processes and its implications for colleges and universities.

Regional accreditation traces its roots to 1885. Today's enterprise is based on decades of experience and refinement, both leading and reflecting the development of American higher education. Many have argued that regional accreditation has provided the conditions under which institutional quality and institutional diversity have flourished because it has allowed institutions to develop unique missions and goals and to be judged on the basis of these rather than some nationally developed institutional model. In each region, standards for all aspects of institutional functioning are carefully developed with input from all campuses within that region. Self-study

processes engage substantial numbers of faculty and staff on each campus, peer review is provided by carefully prepared and reported visitations by external teams, and institutional performance results for each standard are folded together into an overall determination of quality. The performance details are not released publicly, but each institution receives a formal judgment by the Commission and lists of reviewed and accredited institutions are public documents. Accreditation confers a status that is important for institutions, students, parents, the federal government, state governments, and employers.

Another of the lasting outcomes of the Spellings Commission has been the strengthening of accountability measures required by regional accrediting agencies and a stronger emphasis on student learning and achievement as fundamental to judgments of academic and institutional quality. Such emphasis has been demanded by the Council for Higher Education Accreditation (CHEA), the federal oversight and "recognition" agency for regional accrediting bodies. Within the federal recognition is the expectation that accreditation standards and reviews go beyond inputs and processes (for example asking: "Do students have access to learning resources and are they using them?") to focus increasingly on outcomes (for example asking: "How well are students gaining skills of finding, evaluating, and using information?"). Over the past decade, regional accreditation commissions have helped colleges and universities develop trustworthy and useful ways to understand what and how their students are learning and how to use the results for improvement.

The Commonwealth of Massachusetts is part of the New England Association of Schools and Colleges (NEASC) and subject to the standards, guidelines and processes of its Commission on Institutions of Higher Education (CIHE) (NEASC, 2005). In 2008, CIHE introduced its Student Achievement and Success initiative with the following words:

> *The Commission recognizes its dual responsibilities of quality assurance (the public role) and quality improvement (the private role) must be kept in balance.*
>
> *Nowhere is this dual role more in play than in the area commonly summarized by the term 'student assessment.'*
>
> *The Commission's Standards for Accreditation speak to assessment for improving the academic program and services for students.*
>
> *At the same time, the Standards speak to the Commission's public responsibility for ensuring an appropriate level of student success and in making the results used and known – part of its responsibility as being recognized by the federal government as a reliable authority on the quality of education (NEASC, 2008, p. 1)*

In our Phase One report, the Working Group noted that the learning outcomes assessment programs on many public campuses which have been recently reviewed for re-accreditation by NEASC/CIHE include new elements that respond to the new NEASC standards, guidelines and data requirements in the area of student learning assessment. This is a positive development that serves Massachusetts public campuses well. In our Phase One report, we called for "increasing the collaborative spirit between DHE and NEASC with regard to the assessment of student learning." Now that we have examined state-level issues, we feel even more strongly about that recommendation. The press for greater accountability from CHEA, on behalf of the federal government, has led regional accrediting bodies like NEASC to become more accountability- and

results-oriented, especially in the area of learning outcomes assessment, and this has influenced public colleges and universities in Massachusetts.

While the DHE has not historically gathered information on learning outcomes assessment as part of its Performance Measurement System, it has been influenced as well by the assessment movement and state pressures for accountability. The Vision Project declares a public agenda for public higher education in seven outcome areas and commits the system to accountability for progress and leadership. This implies a more public reporting of learning assessment results and public comparisons of student achievements in Massachusetts with those of students in other states. The Working Group sees a challenge in accommodating the legitimate interests and concerns of NEASC and DHE in learning outcomes assessment without increasing the tensions already inherent in the assessment for improvement versus assessment for accountability dynamic.

We applaud the Commissioner and NEASC's CIHE Executive Director and senior leadership for meetings over the past year and their sincere commitments to help Massachusetts campuses strengthen capacities for student learning outcomes and assessment. But faculty members, institutional leaders and some of our own working group members have voiced concern that DHE and CIHE could end up with conflicting expectations and requirements for campuses that could undermine campus efforts in student learning outcomes assessment. The Working Group hopes this will not be the case. We believe that NEASC standards and peer-review-based processes and the new CIHE emphasis in the area of student assessment serve Massachusetts public institutions well and deserve our support. We also recognize and support the importance of the Vision Project agenda for public higher education in Massachusetts and its commitment to measure, compare and report publicly on student learning. We believe that the plan we propose in the next chapter will allow colleges and universities to respond to both.

Judith Eaton, President of the Council for Higher Education Accreditation in a recent *Inside Accreditation* commentary (February 2011), forcefully draws the attention of the higher education community to the problem of "federalizing accreditation." She notes that since the passage of the Higher Education Opportunity Act in 2008, the federal government has increasingly used its power to "recognize" accrediting bodies to reach through them to colleges and universities with an increasing array of regulations. Since their own standing and the flow of federal funds for institutions and students are at stake, accrediting bodies have little choice but to comply with federal regulations. The most recent guidelines for accrediting body compliance reports speak directly to student learning outcomes, general education requirements, academic standards and the like. Eaton notes that this presents difficulty for campuses as they increasingly turn to accrediting bodies for quality assurance and for accountability purposes. The Working Group is concerned about these trends and supports the traditional role of faculty and academic institutions in making academic policy for higher education.

Comparability

The Vision Project sets the bold vision of producing the best-educated citizenry and workforce in the nation and, with respect to student learning, calls for Massachusetts to be a national leader in academic achievements of public college students on campus-level and national assessments of learning. It commits Massachusetts public higher education to measuring student learning and comparing ourselves to other states. The Working Group recognizes the importance of comparability for accountability and the Vision Project. We appreciate the desire for comparability

on the part of the community, investors and key stakeholders external to higher education and recognize that comparative information helps provide the accountability on which public support and needed resources can be gained.

The Working Group also believes that it is very important to understand the limitations and difficulties associated with comparability and to make a firm commitment to "doing it right."

Comparisons in the area of student learning are very complicated and must be undertaken with great care. Their "meaningfulness" or ability to provide valid and reliable information depends heavily on which metrics are selected for reporting and which institutional sets are chosen for comparison. As we noted in our Phase One report, the assessment movement has not agreed upon any single "best" way to measure learning, so it remains important to use multiple measures of learning and triangulate them for better understanding. (Many of these measures will be described in Chapter Two of this report.) Care in the selection of institutions for comparison is important because colleges and universities vary along many important dimensions including, to name a few that have implications for student learning: mission and purpose, characteristics of the student body, urban or rural character, the mix of full-time vs. part-time students and faculty, level of degree offerings, selectivity, and affordability. These differences influence student learning outcomes and the ability to derive meaning from comparisons. "Apples-to-apples" comparisons pay attention to both the appropriateness of the measures and the appropriateness of the comparison set. After a brief review of potential institutional comparison sets, we will make suggestions for two that we believe will be helpful.

Institutional comparisons using peer institutions provide the closest to "ideal" comparisons because peer institutions are selected on the basis of similarity—in student characteristics, union environment, academic offerings, selectivity, control, and the like. They are not new to the Commonwealth. Beginning in 1998, each campus was asked by the Board of Higher Education to select five colleges to be identified as "peer institutions." Every college was then asked to compare itself to each peer institution in four areas: enrollment, program emphasis, high-cost programs and student body characteristics. IPEDS data provided important information for each college to assess, compare, make changes and then reassess. The IPEDS categories included academics, student services, institutional support, plant operations and total expenditures. Peer institution comparisons have become routine in public higher education, although it has been some time since there has been a formal review of the list of peer institutions in the public sector in Massachusetts.

National comparisons of learning outcomes are frequently made through use of nationally standardized tests. The Phase One report discussed the limitations of these instruments as measures of the intended learning outcomes of a given institution and not appropriate as a proxy for academic achievement for the Massachusetts public higher education system. In addition, the national average scores reported for these tests are almost always inappropriate comparisons for any individual institution, since students at all types of institutions may be included in the averages. Comparisons with students from broad groupings of institutions (e.g. all campuses, all publics, all four-year, etc.) cannot provide meaningful information about how well an individual institution with a particular mission and student body is doing in the important work of educating its own students.

State-to-state comparisons. The Vision Project goal for Massachusetts to be a national leader in academic achievement of public higher education students on campus-level and national assessments of learning implies a comparison of the Massachusetts public higher education

system with those of other states. The Working Group believes that state-to-state comparisons are possible with states that have similarly structured systems of public higher education but would be inappropriate with other states. The states selected for comparison would most appropriately have multiple colleges to choose among, both public and private higher education options available, dense population centers—in other words, they would be like Massachusetts. States with single institutions at each segmental level or with populations spread distant from each other are less likely to share the challenges of states with student populations that can easily move from institution to institution.

System-wide, segmental level comparisons. Short of making comparisons for the whole system of public higher education in the Commonwealth with those of other states, some consideration may be given to such comparisons at the segmental level—e.g., comparing the Massachusetts community college system with the community college system in another state. As with other types of comparisons, these will only be meaningful if the structure and mission of the comparison system is similar to that of Massachusetts. For example, some community college systems focus primarily on remediation and transfer; some more on vocational preparation; and some on associate degree completion. Any comparison of learning outcomes from Massachusetts community colleges with those in a state with a dissimilar structure or mission would be meaningless at best and misleading at worst.

A goal for Massachusetts to be the "best" of any comparison group selected is only relevant if the comparison group is appropriate for such comparison. The fact that it is difficult to select cohorts for comparison does not mean that the work is not worth doing, but it is critical that the Commonwealth not settle for selecting easy and expedient comparisons instead of the appropriate ones. The oversimplification of the complex can damage a good assessment system, undermining the very accountability it seeks to provide. Further it will also weaken the support of the faculty, staff and campus leadership charged with implementation of statewide learning outcomes assessment. It is important to engage campuses, faculty and staff in meaningful dialogue about metrics and comparison groups so as to instill confidence in their use.

The WGSLOA believes that peer institutional and peer-state comparisons of student learning outcomes could provide useful information for campuses and the system of public higher education as a whole. We encourage a formal review of peer institutions for each campus, the inclusion of student learning as a category for comparison and the careful selection of appropriate learning metrics, and the continuation of the DHE Linear Trends Report. Peer institutional comparisons offer the best chance for "apples-to-apples" comparisons for each college or university.

Partnering with "Massachusetts-like" states to compare student learning outcomes would provide better information than is now available from efforts to compare the 50 states. It would allow collaboration in the selection of appropriate measures and procedures, for example for shared rubric development and subsequent scoring of course-embedded assessments by faculty in each state. Partnerships could provide each state with meaningful information and, for faculty and institutions involved, could lead to better understanding of how our students are performing.

Transparency

Transparency is an important feature of accountability systems for taxpayer-supported institutions like colleges and universities. In this instance, the concept of transparency implies the willingness

to make visible or public the information that results from learning outcomes assessment. Federal and Massachusetts reporting requirements for public institutions of higher education and the DHE Performance Measurement System provide for a significant level of transparency for many key performance indicators. Retention and graduation rates of students who begin at our institutions as first-time, full-time freshmen and many other data elements are publicly displayed in the Common Data Set which is available on every institution's website. Institutions also create and publish institutional fact books carrying a variety of information related to faculty, staff and student characteristics. Each institution has a story to tell. Such transparency can be very helpful to students and potential students, to parents, and to others who have an interest in the programs and expectations of specific campuses. Using a common format makes it easier to use for external audiences.

Results of learning outcomes assessments are used internally on public campuses in Massachusetts and may appear in self-studies prepared for NEASC but are not routinely made public. The Working Group is not opposed to transparency for assessment results provided that care is taken in the presentation and framing so that the varied constituent groups understand its meaning and its impact on the Commonwealth and the region. Having information publicly accessible but without appropriate explanation can easily lead to the misinterpretation and accidental or blatant misuse of the information. It is critical for us to contextualize any information we report so as to maximize public understanding by clearly communicating what is said by such information and, importantly, what cannot be deduced from the data. Information free from context does not tell the story for the institution, the segment, or the state as a whole. As Peter Ewell noted in our November 2010 conversation, "Part of the problem we face in communicating well to our audiences is that the information they ask for does not actually answer the questions they want to answer." It is important to customize reports to respond to the interests and needs of various audiences.

As this report was very near completion, the Working Group noticed that The National Institute on Learning Outcomes Assessment (NILOA) had posted a new item to the Tool Kit section of its website: " Providing Evidence of Student Learning: A Transparency Framework" (National Institute for Learning Outcomes Assessment, 2011). Described as a work in progress, the transparency framework represents a package of recommendations for how campuses could report learning outcomes information on their websites. It considers issues related to privacy and transparency and recognizes the importance of providing information with sufficient context and explanation for a variety of internal and external audiences as we have discussed in this report. The framework makes suggestions and provides examples in six component areas including: student learning outcomes statements, assessment plans, assessment resources, current assessment activities, evidence of student learning, and use of learning evidence. While the Working Group has not had sufficient time to examine the framework carefully, and the higher education community as a whole has not had time to respond, we think it could be compatible with our proposed statewide plan by offering an additional vehicle for supporting independent campus-based programs, while allowing for commonality in reporting that also serves accountability objectives. We suggest that it be examined as part of the implementation activities described in Chapter Three.

Several states have employed dashboards as a means to display complicated and multi-measured performance data in a manner more accessible for viewers. The Working Group suggests that a dashboard would be a good way of presenting learning outcomes assessment results. Creating a dashboard that displays all the learning metrics selected for the Vision Project

and customizing reports to contextualize the information in different ways to respond to the interests and needs of various audiences would allow for appropriate comparisons with meaningful information, advance the high-quality education provided by public higher education, and allow for quality assurance reporting.

Rankings

Ranking systems in higher education are becoming more and more abundant, and we now see institutions ranked on their selectivity, research expenditures, overall quality, affordability and any number of other things. It has become the case in higher education that everything that is compared is subject to use in some ranking system. But ranking systems can only be as good as the underlying selection of metrics and comparison groups allow, and can only be meaningful if the results are accurately reported with sufficient context for understanding. Without careful attention to the basics, rankings can as often misinform as inform.

Malcolm Gladwell (2011) in a recent article in *The New Yorker* shows how small modifications in the variables (metrics), or in how variables are measured, would have produced vastly different rankings of colleges and universities in the very influential rankings provided annually by *U.S. News & World Report.* He shows how reputational prejudices and implicit ideological choices influence the rankings. He notes how the failure to include price data in the ranking system deprives consumers of information about affordability and how the choice to weight selectivity more highly than efficacy (e.g., better graduation rate performance than would be expected given the socioeconomic and test score performance of incoming students) produces rankings skewed toward highly selective private institutions. He concludes: "Who comes out on top, in any ranking system, is really about who is doing the ranking" (p. 75).

Failure to understand the metrics used for comparison can contribute to rankings that can misinform or intentionally mislead the public. The metrics that are typically reported on a dashboard are based on averages as representative of institutional performance. The variability within an institution on metrics, however, may be as great as, or greater than, the variability between institutions. Given this, it will be critical to consider any variations between institutions with care. Careful selection of appropriate metrics and comparison groups facilitate better cross-institutional comparisons.

It is also important to recognize that rankings often fail to take into account differences that we know exist and we know to be important. For example: public higher education institutions often focus more on affordability and accessibility than do private institutions. State rankings that fail to account for differences in the mix of public and private institutions in the state and for differences in institutional mission can be seriously misleading. Accurate institutional and state peer groups could help each institution provide its own information straightforwardly and with appropriate context.

Gladwell notes that ranking systems that try to be heterogeneous by working with many different types of entities or institutions, like colleges and universities of all types and sizes, **and** that try to be comprehensive as well by taking many different variables into account, like the seven weighted variables in the *U.S. News* college rankings, present real difficulties. "But it's an act of real audacity when a ranking system tries to be comprehensive *and* heterogeneous—which is the first thing to keep in mind in any consideration of *U.S. News & World Report's* annual 'Best

Colleges' guide" (Gladwell, 2011, p. 70). He shows clearly how small changes in the weightings of the variables produce very different lists of "best" institutions.

Many ranking systems for states already have Massachusetts at or near the top. The question to wrestle with is how meaningful the data are on which institutions are being ranked in these systems. We have an opportunity with the Vision Project to select meaningful metrics and understand better the variables that must be controlled for when selecting public institutions of higher education with which we will be compared and ranked.

The WGSLOA shares with many faculty members a concern with the use of the term "best" in the Vision Project and thinks that the use of "best" exacerbates the tensions between the faculty/institutional perspective and the state perspective. At the Working Group's meeting with faculty in January 2011 one faculty member summed up the issue with her comment to us: "What will they do with best? If you're best does that mean that you obviously have enough resources and don't need more? If you are far from best will resources flow to you or be taken away? Why would peer institutions or states share information for comparison with us if we are setting out to prove that we are best?" These questions have been important for the consideration of this Working Group and are important for the Vision Project as a whole.

The WGSLOA recognizes that rankings are a reality in the public arena in which we work. As a result we believe it is especially important for us to help develop reporting systems for learning outcomes assessment with meaningful information about our institutions and the system as a whole. The Working Group believes that Massachusetts can be a leader in this effort and, in doing so, will make an important contribution for higher education.

Patterns and Practices in Other States

By late 2009 when the Working Group began its work, there was a growing consensus across the United States that there needed to be more accountability in public higher education, particularly with regard to student learning outcomes. A number of states had already begun to measure and publicly report on student learning in their public institutions of higher education. But 38 states, including Massachusetts, did not report data on student learning in any systematic form (Aldeman & Carey, 2009). Across the states, a variety of different approaches have been adopted that attempt to comprehensively assess student learning at public institutions. These efforts range from statewide mandated standardized tests to common required state learning outcomes assessed in different ways on different campuses.

A variety of standardized tests are used in different states including CAAP (Collegiate Assessment of Academic Proficiency), MAPP (Measure of Academic Proficiency and Progress), CLA (Collegiate Learning Assessment), the California Critical Thinking Skills Test, and C-Base (College Basic Academic Subjects Exam). A number of states have also adopted or are piloting a set of common learning outcomes and associated rubrics through the program LEAP (Liberal Education and America's Promise) initiated by the AAC&U (Association of American Colleges and Universities). Six states (California, North Dakota, Oregon, Utah, Virginia and Wisconsin) have been designated LEAP States, and another dozen have expressed interest in being so designated. Additionally, Arizona, Connecticut, Iowa, Kentucky, North Carolina, North Dakota, and South Dakota have initiated efforts to develop state accountability systems.

A number of studies and reports attempt to describe, categorize or grade the 50 states in the area of higher education (Callan, 2006). From 2000 to 2008, the National Center for Public Policy and Higher Education published "Measuring Up," a biennial state-based report card on higher education covering a variety of outcome areas. Work for "Measuring Up" was an important activity of the National Center for Higher Education Management Systems and Peter Ewell (Ewell, 2008). For obvious reasons, student learning was always considered an important factor for comparison, but for years "Measuring Up" reports gave most states an "incomplete" in this category because there was not enough information on student learning available for satisfactory comparison.

Aldeman and Carey's 2009 report for Education Sector assessed state higher education accountability systems and gave 38 states a designation of "needs improvement" in the area of student learning. They gave 10 states (California, Florida, Georgia, Hawaii, Kansas, Maryland, Minnesota, Missouri, New York and Virginia) a designation of "in progress" for a varied set of initiatives and practices. Only four states (South Dakota, Tennessee, Texas and Wyoming) earned the designation of "Best Practice" for having the following features in their student learning outcomes category: "the inclusion of entire systems of institutions, the ability to make comparisons across institutions, and some type of objective assessment" (Aldeman & Carey, 2009, p. 2).

In a recent *NILOA* Paper, "Connecting State Policies on Assessment with Institutional Assessment Activity," Ewell, Jankowski and Provesis used the results of the same national *NILOA* survey of provosts used in our Phase One work and an *NCHEMS* study on state policies to compare the practices of eight "assessment-active" states with those of less active states. While they cannot be fairly described as "model" or "best practices" states, the convergence of opinion on the major players in student learning outcomes assessment is worth noting. The states included: Georgia, Kentucky, Minnesota, Oklahoma, Rhode Island, South Dakota, Tennessee and West Virginia (Ewell, Jankowski & Provesis, 2010).

States also vary in their approach to reporting their student learning data. A number of states with a single governing board for all four-year institutions or for all universities have adopted the Voluntary System of Accountability (VSA). In these systems, campuses in VSA use one of three standardized tests (CAAP, MAPP or CLA) and report the results, along with other measurements of student success, in an online College Portrait. The intent is to provide for comparability. The VSA has been criticized, however, for failing to help consumers compare data across institutions because of incomplete data, differences in assessments, and differences in the ways the data are presented (Aldeman & Kelly, 2010).

Some states have chosen to present their own student learning and success outcomes through online "dashboards" that summarize data for their institutions of public higher education. In addition to providing data on student learning, some states provide data on student engagement from the National Survey of Student Engagement (NSSE) and the Community College Survey of Student Engagement (CCSSE). The small number of states that have used learning assessments for many years often report the results as part of Performance Assessment Systems (e.g., Texas, Tennessee) while Maryland prepares narrative reports on the assessment program of each public campus.

The fact that many states and campuses do not use standardized tests, the diversity of standardized tests in use, even within individual states, and the policy of the VSA to allow institutions to select either the CAAP, MAPP or CLA makes it currently not possible to make the

kind of "among-all-states" comparisons on national assessments of learning suggested by the Vision Project.

While the VSA suggests that the CAAP, MAPP and CLA provide comparable data, there are important differences between and even within these instruments. For instance, the CAAP can be administered in individual modules, the MAPP has both abbreviated and standard forms, and the CLA is a very different format of assessment involving open response rather than multiple choice. Among the 14 "best practice" or "in progress" states in learning outcomes assessment identified by Aldeman and Carey (2009), only five used a single standardized test across all of their institutions of public higher education, and only two of those—Minnesota and South Dakota— used one of the three instruments approved for use in the VSA . While both Minnesota and South Dakota used CAAP, Florida, Georgia and Hawaii each use instruments that are not widely used in any of the other states. CAAP is also used in Kentucky, and CLA is used in the West Virginia system. Based on this variability of standardized testing, to effectively compare itself with other states, Massachusetts may be best served by allowing the same flexibility in the use of standardized tests in our system.

In contrast to the diversity of standardized assessments across states, there is arguably less diversity in the application of common outcomes and rubrics. While many states allow campuses a great deal of autonomy in establishing their own learning outcomes and means of assessing those outcomes, the existence of a single, widely recognized national set of outcomes and rubrics (LEAP) means that an increasing number of states are moving to adopt these outcomes. California's State University system, which also uses the CLA as its standardized test and has adopted the VSA, is part of LEAP. In addition, North Dakota, Oregon, Utah, Virginia and Wisconsin are all LEAP states, while Kentucky and Texas are both discussing the possibility of adopting the LEAP outcomes. While LEAP does not yet have the body of data for cross-institutional comparisons that is available through standardized tests, its adoption across a number of statewide systems provides promise for its future use in between-state comparisons. It should also be noted that the kind of embedded assessments required for use in LEAP are widely considered among the best ways to measure student learning.

In terms of a method for reporting student learning outcomes the VSA is similar to LEAP in lacking any common alternative approach. Therefore, the VSA is fairly widely utilized across many states. Thirty-one of 33 Texas public universities and seven of eight of the University of Louisiana System campuses use the VSA. In addition, all of the California State University, University of North Carolina system, Pennsylvania State System of Higher Education (PASSHE), and University of Missouri campuses use the VSA, as do all of the public Ohio, Kansas, Iowa and University of Wisconsin System universities. As noted above, the VSA has been criticized for the ways in which data is collected and presented because it falls short as a tool to compare individual campuses. This critique suggests that it may not be a valuable tool for comparing the Massachusetts institutions of higher education to those in other states. Furthermore, because unlike LEAP it lacks common learning outcomes and only defines learning in terms of institution-specific outcomes and performance on standardized tests, it is not critical for a Massachusetts institution to participate in VSA in order to compare data from that institution to others in the VSA system. Instead, such comparisons would depend primarily on the selection of standardized tests utilized in Massachusetts. On the other hand, while there may not be system-wide advantages to adopting VSA, there may be accreditation advantages to individual campuses.

In summary, over four years after the Spellings Commission report, there remains a great deal of variability in the assessment of student learning among institutions of higher education and

among states. The wide variation in assessment approaches and in the ways of reporting assessment results can be explained in part by the wide variation in structure and governance for higher education and by the different influences of regional accrediting bodies whose expectations constrain and guide institutions. States in which most or all campuses are part of a single system under a single governing board more often have unitary practices and single reports while statewide coordinating agencies are characterized by far greater diversity of practice and approach. Different collective bargaining arrangements with faculty also introduce considerable variation in state policy and practice, as do differences in student body characteristics and program mix among institutions. Such diversity is often considered a hallmark of the American higher education system and a source of its strength, but it also makes it more difficult to identify model states and best practices. In our examination of other states, we were able to identify states which are more active in the area of student learning assessment, but we did not find clear evidence of "model" or "best practice" states that should guide our thinking in Massachusetts.

The variability among states makes it difficult to select appropriate ways to compare Massachusetts to other states, but the Working Group finds that the LEAP initiative holds the most promise for providing opportunities for collecting and comparing data on student learning across different states and institutions. We also conclude that Massachusetts should develop an approach that will work for the unique combination of institutions and governance patterns in the Commonwealth and, as we noted in the Phase One report, that builds on the assessment programs on campuses.

Moving Forward

Betterment is a perpetual labor. The world is chaotic, disorganized, and vexing, and medicine (education) is nowhere spared that reality. To complicate matters, we in medicine (education) are also only humans ourselves. We are distractible, weak, and given to our own concerns. Yet still, to live as a doctor (educator) is to live so that one's life is bound up in others' and in science (teaching and learning) and in the messy, complicated connection between the two. It is to live a life of responsibility. The question, then, is not whether one accepts the responsibility. Just by doing this work, one has. The question is, having accepted the responsibility, how one does such work well. (Gawande, 2007, .p 9)

The assessment of student learning improves learning. The faculty and staff within the colleges and universities of the Commonwealth of Massachusetts have a longstanding and evolving commitment to student learning and the measurement of student progress. Just as there are good questions about the level of accountability that exists at the core of the assessment reform movement, there are also good questions at the core of concern found within the academy with regard to the oversimplification of the very complex process of learning. It will only be as a result of the difficult and ongoing conversations at the campuses, in regions and throughout the system that reform will engage the passions and innovative thinking of faculty and staff. That engagement will elevate the confidence and trust of the key stakeholders of public higher education in local, campus-based decisions and in the accreditation system, and will lead to increased commitment to public higher education. We believe that this will ultimately lift and improve student learning.

The Vision Project provides the opportunity to create a system of assessment that has intention to develop a culture of inquiry and improvement within the public higher education system; support existing efforts at each college and university; support faculty and staff in the

development of their own learning about assessment; increase the system's accountability to future students, their families and the taxpayer; increase the accountability at each college and university to their own faculty, staff and students; and increase the confidence and commitment of key stakeholders and policy makers in the state public higher education system.

Chapter Two
A Plan for Statewide Learning Outcomes Assessment

If we are to accomplish the goal set forth in the Vision Project of becoming a national leader in "academic achievements on campus-level and national assessments of student learning," then our system of 29 public colleges and universities[1] must have a plan for establishing, measuring, and reporting learning outcomes that allows for comparability between Massachusetts and other states, even as it addresses the most important feature of any campus-level assessment activity: the improvement of student learning.

In Chapter One, drawing on the work of Peter Ewell, we presented the two paradigms that result from the different purposes of assessment and endorsed his four principles for resolving the tensions that result from the different paradigms. We also discussed the issues of comparability and transparency and the difficulties both can present for state systems and for campuses. In this chapter we begin by discussing the challenges these issues present for the development of a statewide plan for learning outcomes assessment. With the Ewell principles and our challenges as a starting point, we develop a model for statewide learning outcomes assessment. The model provides the conceptual underpinning for a plan for statewide learning outcomes assessment that engages all 28 undergraduate campuses and the Department of Higher Education.

We believe that this plan reflects best thinking and best practice in learning outcomes assessment and responds to the needs of the Vision Project. It would represent the collective effort and responsibility of the campuses and the Department of Higher Education and our collective willingness to be held accountable.

The Challenges of a Plan for Statewide Learning Outcomes Assessment

Challenge 1: Improvement & Accountability

> *The Challenge: Any statewide system of learning outcomes and assessment should somehow balance the unique, discipline-specific feedback that faculty need in order to make actual improvements in teaching and learning with the broader, publicly communicated and comparable results of more general assessments.*

At most colleges the assessment of student learning began and continues to be conducted for one of two reasons: **improvement** or **accountability**.

Assessment for **improvement** is about improving teaching and learning. It typically involves faculty members gathering evidence about how well students are accomplishing the intended learning outcomes of a course, or a program of study, and then using the evidence to improve

---

[1] There are 29 public higher education campuses in Massachusetts, 28 of which offer undergraduate programs. This report uses 29 when referring to the system in general terms and 28 when referring specifically to the undergraduate campuses.

instruction and institutional policies. The nature of the evidence gathered and the way it is analyzed and applied are usually specific and unique to the instructor, the program, the academic discipline, and the curriculum of a particular college. This type of assessment is mostly intended for, and best understood by, *internal* audiences (the faculty and staff of the college).

Assessment for **accountability**, on the other hand, is about demonstrating to the public that a college is effectively using its resources to help students learn (and, often, to contribute to society and the workforce). The nature of the evidence gathered and the way it is analyzed and applied are usually more general, perhaps standardized, related to broad skills, and more easily communicated to non-academics. This type of assessment is mostly intended for *external* audiences (potential students and their families, the public at large, governors and legislators, etc.)

In addition to noting that these two types of assessment represent distinct paradigms as discussed in Chapter One, Peter Ewell also notes that these two purposes of assessment may contain "opposite incentives." That is, when assessing for **accountability**, institutions are incentivized to "look as good as possible, regardless of the underlying performance," while "discovering deficiencies is one of the major objectives of assessment for **improvement**." Balancing improvement and accountability also requires resolving contradictory incentive structures (Ewell, 2008).

Challenge 2: Privacy & Transparency

> *The Challenge: Any statewide system of learning outcomes assessment must find ways to secure what must be secured, such as individual student records, and provide safe environments in which to discuss gaps, problems, and challenges to learning, while at the same time appropriately engaging and informing the public.*

By its nature, assessment for the purposes of improving teaching and learning involves the work of individual students, the performance of individual faculty and staff, and the considerable experimentation, trial and error of everyone involved. A safe environment in which problems and solutions can be analyzed is necessary, and there are compelling reasons to maintain some **privacy** around select assessment practices and products.

At the same time, in order to be publicly accountable and provide information that is comparable, there must be some amount of **transparency** in the establishment of benchmarks, the creation of assessment goals and tools, and the analysis and presentation of data that is gathered.

Challenge 3: Uniqueness & Comparability

> *The Challenge: Any statewide system of learning outcomes and assessment must balance the strengths of our colleges' and universities' unique missions and ways of delivering education with the very real need to compare the accomplishments of higher education in Massachusetts with other states.*

One of the great strengths of American public higher education, and of the Massachusetts higher education system, is the relative autonomy and **uniqueness** of each of our colleges and universities. Each institution is free to develop and deliver academic programs, conduct research, and invent products and services according to its unique mission, guided and governed by its own policies and procedures. Faculty members have academic freedom and the expectation of participating in the governance of their institutions. Accrediting and governmental agencies exercise limited and specified oversight. The result is a national "marketplace of ideas" in which colleges freely compete for students, faculty, resources, partnerships, and prestige.

At the same time, the very uniqueness and freedom of colleges and universities make **comparability** across campuses, and from state to state, extremely difficult. In an era of diminishing resources, and diminishing trust, for public higher education, we are increasingly called upon to provide information that allows the public to compare not only the offerings, but also the cost, the quality, and the outcomes of learning that occur at our schools.

A Model for Statewide Learning Outcomes Assessment

If the colleges and universities in Massachusetts are to devise a system for assessing the learning outcomes of their undergraduates in a way that allows for comparability, transparency, and accountability, we must agree on some of the *qualities* of an undergraduate education that we all expect our students to possess. At the same time, those qualities we agree on must allow us to preserve the unique missions of individual colleges, appropriate measures of privacy around assessment work, and an ability to actually improve teaching and learning with the results that we find.

Additionally, research on sound assessment practice strongly suggests that no single instrument or approach to assessing learning can meet all of the challenges described above. The most effective systemic models of assessment offer *multiple measures* of student learning in a "triangulation" approach that includes indirect assessments such as surveys, direct assessments like tests, and embedded assessments such as classroom assignments.

Our model for Massachusetts public higher education is presented in Figure 1 on page 24. It illustrates the use of the "LEAP Essential Learning Outcomes" as a flexible framework for statewide learning outcomes assessment (Association of American Colleges and Universities, 2007). Within this framework, assessment occurs in three ways, as depicted on an inverted pyramid. At the "point" of the pyramid, individual faculty and their students are engaged in "embedded assessment" of learning outcomes, through classroom assignments, research papers, lab demonstrations, capstone projects, and other artifacts of actual coursework. Some are engaged in embedded assessment as a result of specialized program accreditation. Across the middle of the pyramid, larger numbers of students are involved in "direct assessment" as they take professional licensure exams or admissions exams for graduate or professional schools, or standardized tests selected by their programs or campuses. At the inverted "base" of the pyramid, undergraduates across the campuses of all of Massachusetts' public colleges and universities are involved in "indirect assessment" as they participate in either the Community College Survey of Student Engagement (at two-year colleges) or the National Survey of Student Engagement (at four-year colleges and universities). Faculty and staff use information from the surveys to improve the curriculum and the cocurriculum in support of student learning.
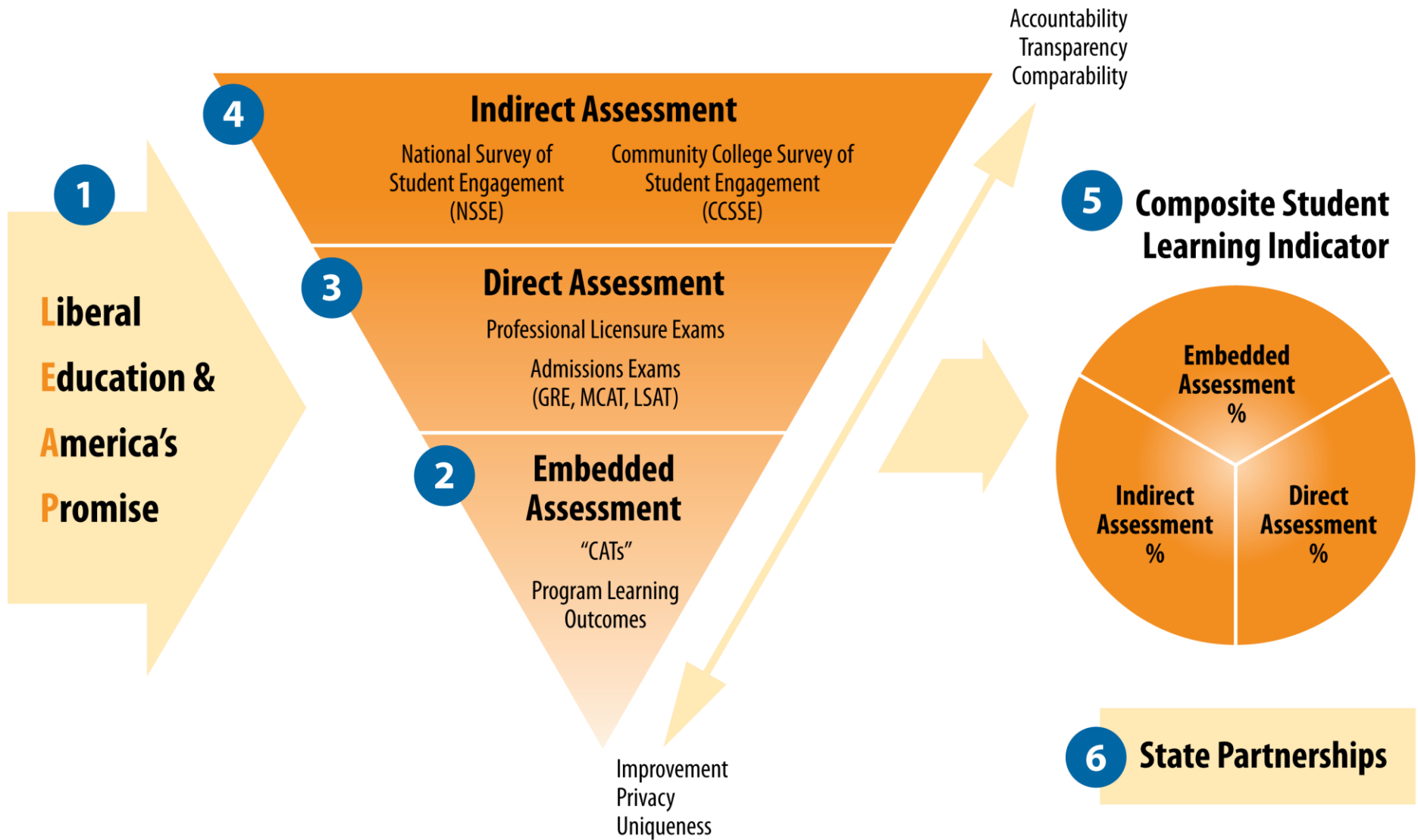
Each of the items in the pyramid exists along a continuum between "Improvement and Accountability," "Privacy and Transparency," and "Uniqueness and Comparability," and each can be publicly reported through a dashboard, profile or other means. Together they offer the Commonwealth a robust campus-based system of collegiate learning outcomes assessment.

Additional features in the model involve system-level work. Results from campus assessments can be reported at the system level in a multiple measures report presentation or dashboard. They can also be combined into a composite learning indicator. Another feature, state partnerships, could be formed and used for comparing learning outcomes results across states. All system-level work adds capability for **statewide accountability, transparency and comparability.**

Taken together the LEAP framework, the campus-based multiple measures of the pyramid, and the system-level features provide a conceptual model that supports a Plan for Statewide Learning Outcomes Assessment.

Figure 1

A Model for Statewide Learning Outcomes Assessment in Massachusetts

Accountability
Transparency
Comparability

1

Liberal

Education &

America's

Promise

4
Indirect Assessment
National Survey of Student Engagement (NSSE)

Community College Survey of Student Engagement (CCSSE)

3
Direct Assessment
Professional Licensure Exams

Admissions Exams (GRE, MCAT, LSAT)

2
Embedded Assessment
"CATs"

Program Learning Outcomes

5 **Composite Student Learning Indicator**

Embedded Assessment %

Indirect Assessment %

Direct Assessment %

6 **State Partnerships**

Improvement
Privacy
Uniqueness

A Plan for Statewide Learning Outcomes Assessment

We have turned the major features in our model into components of a Plan for Statewide Learning Outcomes Assessment and will further elaborate the components in this chapter (following the numbers in the model). Our discussion of the components is based on information gathered from a national survey of provosts (Kuh & Ikenberry, 2009), the DHE survey of campus provosts undertaken during our Phase One work, from literature and best practice, from our examination of state-level issues, and from our conceptual model. It is important to note that we have only broadly sketched the components. More detailed development of the Plan and its components needs to involve a larger group of faculty and staff, assessment and institutional research professionals, academic vice presidents and provosts, representatives from the Department of Higher Education, and campus presidents than was represented by our Working Group because ultimately this is a plan **of the public campuses** in the system, not **for the public campuses** in the system.

1. Liberal Education and America's Promise (LEAP)

Liberal Education and America's Promise (LEAP), created by the Association of American Colleges and Universities (AAC&U), is a national initiative advocating the importance of a "twenty-first century liberal education" that provides a framework for the statewide plan for learning outcomes assessment (Association of American Colleges and Universities, 2007). The LEAP "Essential Learning Outcomes" speak directly to the qualities that we want our undergraduate students to possess. The list of LEAP learning outcomes appears in Appendix B but some of the qualities include:

- Knowledge of Human Cultures and the Physical and Natural World (Science, Math, Social Science, Humanities, History, Languages, and the Arts)

- Intellectual and Practical Skills (Critical and Creative Thinking, Written and Oral Communication, Quantitative Literacy, Information Literacy, Teamwork and Problem Solving)

- Personal and Social Responsibility (Civic Knowledge and Engagement, Intercultural Knowledge and Competence, Ethical Reasoning and Action)

- Integrative Learning (Advanced Accomplishment Across General and Specialized Studies)

As noted in our Phase One report, considerable similarities already exist among the various institutional learning outcomes in place at Massachusetts public colleges and universities and the

"Essential Learning Outcomes" expressed by LEAP. If adopted, the LEAP framework can provide us with necessary *comparability* across institutions, and with other "LEAP States."

As the LEAP Essential Learning Outcomes have been developed and communicated more broadly, and as more colleges and universities have adopted the framework for their institutional measurements of learning outcomes, many additional LEAP-related features have emerged that hold considerable promise for both improving and publicly reporting student learning:

a. Rubrics for faculty to use when assessing student achievement in 15 different liberal learning outcomes have been developed as part of AAC&U's Valid Assessment of Learning in Undergraduate Education (VALUE) initiative (Rhodes, 2010). The rubrics, which were designed by teams of faculty from across the country, may be used as they currently exist, or they may be adapted to fit the needs of specific states or campuses. Appendix C provides brief definitions and a listing of VALUE Rubrics.

b. A number of "High Impact Practices" have been identified that have proven to be beneficial for college students from many different backgrounds (Kuh, 2008). The list includes practices such as Learning Communities, Service Learning, Study Abroad, Research with Faculty and Senior Culminating Experiences. High Impact Practices are becoming increasingly common at public colleges and universities in Massachusetts. Several years ago, AAC&U formed a partnership with California, Oregon and Wisconsin and with several campuses in each state and they worked together on a "Give Students a Compass" project to improve general education and

> **IN THE WORDS OF THE FACULTY**
>
> *"Good assessment takes a very long time and comes from the faculty."*

ensure that all types of learners from all types of institutions have opportunities for a quality liberal education and can achieve essential learning outcomes. Work with Compass has now been followed with a "Road Map" project with additional states and campuses and a focus on ensuring that traditionally underserved students have access to, and take advantage of, high-impact practices (Association of American Colleges and Universities, 2011). The Road Map work is important because it brings together the curriculum and the cocurriculum as well as faculty and staff in support of student learning and student success and supports the goal to achieve comparable outcomes among different ethnic/racial, economic and gender groups.

c. LEAP connects to faculty and staff, campuses, and state systems of public higher education (Association of American Colleges and Universities, 2008). As noted in Chapter One, several states have begun to build LEAP features into statewide policies and assessment programs and many states, including Massachusetts, are considering affiliation with AAC&U as LEAP States. This creates the potential for developing a partnership among LEAP states for developing approaches rooted in formative assessment and allowing for comparability across states. The LEAP Campus Action Network comprises more than 300 colleges and universities and brings together academic and student affairs leaders to share best practices in undergraduate education and strengthen educational achievement.

d. LEAP also connects to K-12 education, workplace and business communities and other organizations. AAC&U advocates Principles of Excellence and Learning Outcomes as frameworks that support pathways from school to college (Association of American

Colleges and Universities, 2011). AAC&U has joined with the Council on Higher Education Accreditation (Association of American Colleges and Universities & The Council for Higher Education Accreditations, 2008) and many national foundations including Carnegie, Met-Life and Lumina in support of policies, programs and projects that emphasize high-quality liberal education for all students. It has worked with business leaders and conducted several studies of employer's views on student learning. The most recent study found strong support for the broad liberal education advocated by LEAP, for LEAP essential learning outcomes and for versatile knowledge and skills that can be applied in "real-world" settings (Hart Research Associated, 2009). The study echoes many of the priorities and recommendations of the CAGUE Report prepared at the request of the Board of Higher Education in 2009 and reviewed as part of Phase One.

Importantly, though, LEAP does not call for a "one-size-fits-all" curriculum, but instead acknowledges that, "The recommended learning outcomes can and should be achieved through many different programs of study and in all collegiate institutions, including colleges, community colleges, and technical institutes, and universities, both public and private (Association of American Colleges and Universities, 2007, p. 2). This preserves the necessary *uniqueness* of Massachusetts' colleges and universities.

2. Embedded Assessment and Specialized Program Accreditation

Individual faculty and academic programs at most colleges have been involved in embedded assessment techniques for many years now. What began as "classroom assessment techniques" designed to give instructors instant feedback on student learning during a classroom lesson has evolved into a sophisticated array of tools and techniques for designing classroom assignments, and gathering and analyzing student work in ways that can lead directly to the improvement of instruction and institutional policies. Embedded assessments are the most tailored to the individual needs of faculty and students. They often require some measure of **privacy** and are generally considered the most useful form of assessment for actual improvement of teaching and learning. Because embedded assessments are directly linked to student work, they are often considered the most authentic measure of student learning outcomes

One of the benefits of embedded assessment is that the same measures used to evaluate student performance in a course can also be used to assess the extent to which students are achieving the stated learning outcome of an academic program or the campus as a whole. The LEAP framework, and particularly the LEAP VALUE rubrics, offer an opportunity for Massachusetts public colleges and universities to combine the best features of individual faculty and program embedded assessment techniques with a system for comparing assessment results within and across campuses and even across states. The adoption of student electronic portfolios (online repositories of student work) is one mechanism for facilitating this work. If we adopt the LEAP framework and incorporate the use of the LEAP VALUE rubrics into institutional assessment plans—even with variations according to the mission and needs of individual colleges and universities—we can retain the necessary *privacy* of student information and make genuine *improvements* to student learning, even as we provide for *comparability* and greater *transparency* of our student achievement data. Within the LEAP framework for Student Learning Outcome Assessment, embedded assessment is where the "rubber meets the road," providing an authentic, direct measure of student learning outcomes upon which program improvement and accountability can be ascertained.

Embedded Assessment for Academic Programs and Campuses

While traditional classroom assessment such as homework, quizzes and exams are most often used to measure students' understanding of course content, embedded assessments are often used specifically to determine whether or not a program is producing graduates with the requisite knowledge, skills and attitudes to succeed in the workplace. The embedded assessment process at the academic program level involves a group of faculty agreeing on learning outcomes and then embedding specific measures within regular course assignments to determine how well students are acquiring the necessary knowledge or skill and whether or not the intended outcomes are being achieved.

For example, faculty in an Early Childhood Education program might look at their students' ability to "identify various resources and community agencies that support families with young children." In order to measure student achievement on this learning outcome, they might choose to build this ability into a research paper or community service learning project, and include the outcome on a rubric that is used to grade the assignment. Then, at the end of the semester or academic year, the faculty could compare student performance on this specific outcome and determine whether and how changes should be made to what is being taught, how it is being delivered, textbook materials, the instructions for the assignment, and other features of the course.

In many academic programs an extra step is taken to convert broadly expressed learning outcomes (often referred to as goals) into more specific learning objectives for each course to more precisely capture student learning and skill development. Student learning objectives represent the cognitive (i.e., what graduates are expected to know), behavioral (i.e., what graduates are expected to be able to do), and affective (i.e., what graduates are expected to value). The results of embedded measures of learning objectives over a number of courses can be added together to provide demonstrable evidence of students' successful achievement of the learning outcomes for the academic program.

Embedded assessment can include portfolios that contain collections of student work used to demonstrate student growth and achievement in a particular area of study, as well as capstone courses that provide the opportunity to assess the extent to which students are able to synthesize the knowledge and skills they have acquired in a major field of study.

The same process used with academic programs can be used for learning outcomes that are developed for campus-wide application, with larger faculty committees deciding on the outcomes desired for all students regardless of academic major. Faculty members teaching core, general education or capstone courses embed assessment items in the courses, and the results are compiled to determine if campus-wide outcomes have been achieved.

Specialized Program Accreditation

Another benefit of including embedded assessment in our learning outcomes model is that many faculty members are already deeply involved in the practice as a result of work they perform for specialized program accreditation. Specialized program accreditation involves external associations, representing the "field," establishing standards, criteria and processes for formal program accreditation in a manner analogous to processes used by regional accrediting associations for review of campuses [e.g., the Accreditation Board for Engineering and Technology (ABET), the Association to Advance Collegiate Schools of Business (AACSB), the National Association of Industrial Technology (NAIT) and many others]. The accrediting body

assures the quality of the program for external employers and/or regulatory agencies. Embedded assessment is often used in self-studies or reports prepared for these program-specific accrediting agencies.

An example of the student learning goals developed by the Mechanical Engineering Technology (MET) program faculty at Springfield Technical Community College (STCC) in Springfield, MA, for compliance with ABET's Criteria for Engineering Technology programs is shown below in Table 1. Continuing the example, Table 2 provides a list of the student learning objectives for a specific course in the STCC MET program that were developed and agreed upon by MET faculty and stakeholders in support of the student learning goals listed in Table 1.

Table 2 – Example of Program-Level Student Learning Goals

Upon graduation from the MET program, the student is expected to be able to demonstrate:

1. Competence in the knowledge and skills of a mechanical engineering technologist

2. An ability to apply current knowledge and skills and adapt to emerging applications in technology

3. An ability to design and conduct experiments, as well as apply results to improve processes

4. An ability to apply creativity to the design of parts and processes

5. An ability to function effectively on single-discipline and multi-disciplinary teams

6. An ability to identify and solve technical problems

7. An ability to write, speak and listen effectively

8. Recognition of the need for, and an ability to engage in lifelong learning

9. Knowledge of professional standards and ethics and an ability to apply them to his/her work

10. Knowledge of the impact of engineering solutions in a global and societal context

11. Recognition of the need for and an ability to apply timeliness, a commitment to quality and continuous improvement to his/her work

Table 3 – Example of Course-Level Student Learning Objectives

Upon completion of this course, students will be able to:

1. Use 2D AutoCAD to accurately describe mechanical parts, electrical/electronic circuits, and create electrical schematics and wiring diagrams.

2. Demonstrate the ability to utilize technology by accessing course materials and submitting work via BlackboardÒ course management system

3. Communicate drawing information effectively

4. Incorporate appropriate standards for engineering drawing generation

5. Demonstrate timeliness in the completion and submittal of work

6. Demonstrate proficiency with AutoCAD 2D fundamentals: object drawing and editing, template management, dimensioning techniques and dimension styles, layer set up (color, line type and line weight) and use, and drawing accuracy techniques).

ABET and other specialized accrediting bodies require explicit evidence to show how student learning objectives are aligned with both program-level student learning goals and the published learning outcomes and mission of the institution. This alignment is often achieved using a course "crosswalk," a matrix or table that shows *in which course* and *to what extent* each of the stated student learning goals is addressed. The course crosswalk represents a simple and convenient method for identifying potential gaps in a program's curriculum. During an accreditation visit, programs are required to present samples of student work (e.g., homework, quizzes, tests, reports, portfolios, etc.) and show specifically how each of the student learning outcomes and goals are being achieved. Accreditation also requires that academic programs have in place a documented process by which the student learning goals and objectives are developed and periodically updated to reflect the needs of the constituencies served by the program, and have a curriculum that clearly articulates a path for students to achieve the programs learning goals.

3. Direct Assessment

Direct assessments of student learning use actual student work to provide evidence of the degree to which students have attained specified learning outcomes. Faculty members routinely engage in direct assessment when they review and provide feedback to students on their work through the use of the embedded assessments previously described.

There are, however, other forms of direct assessment with the potential to provide more comprehensive feedback to academic programs, to individual colleges and universities, and to the entire public higher education system. Graduate school admission test scores and professional licensure examination pass rates can provide important evidence to assess the quality of preparation of undergraduate degree programs in Massachusetts and effectively balance all six challenges identified in the model:

- They provide specific information that can be used by individual faculty and institutions for the purpose of *improvement;* even as their public nature already holds institutions *accountable* through associated accreditation processes.

- Because the results of licensure exams may be easily gathered and aggregated for public distribution, they can address the need for *transparency;* even as individual student or institutional results are appropriately guarded, thus providing appropriate levels of *privacy.*

- As the licensure exams are already built into student and institutional expectations, the *uniqueness* of each college or university is preserved by the manner in which the academic programs are being offered; while *comparability* is provided through the standardized nature of the exams and the ease of reporting results.

The results of student achievement on graduate admissions and licensure examinations are often already being gathered and presented to accrediting agencies, used in institutional program reviews, reported to the public via college websites and other publications, and included in

institutional and system accountability frameworks. The challenge is to find a means of taking the next step and reporting them in aggregated format as a state system in a manner that is accessible, understandable and meaningful to the general public.

The use of this data as a student learning outcomes assessment measure must be handled carefully in light of the comparability issues raised by the varying range among institutions of entering student educational attainment and academic skill levels. Based on entering student abilities, what would be the expected average institutional test scores and pass rates? How do the actual scores differ from these expectations? How can data be collected and reported for appropriate sets of peer institutions?

Standardized Tests to Measure Student Learning Outcomes
A number of standardized tests to measure a small, specified set of student learning outcomes are in use by colleges and universities across the country and on several public campuses in Massachusetts. Examples of the primary types of student-learning outcomes measured by the various tests currently on the market include reading, writing, mathematics, critical thinking and problem-solving skills. Although there is no consensus among assessment practitioners about the most effective test instrument, the following three tests have emerged as prominent in the field: (1) the Collegiate Learning Assessment (CLA), (2) the Collegiate Assessment of Academic Proficiency (CAAP) and (3) the Measure of Academic Proficiency and Progress (MAPP) test.

As noted in Chapter One, the Voluntary System of Accountability (McPherson & Shulenberger, 2006) and a number of state systems require colleges and universities to use one of these three tests (CLA, CAAP or MAPP) in an attempt to collect and report comparable student learning outcomes assessment data. A Test Validity Study, undertaken under the auspices of the Fund for the Improvement of Postsecondary Education (FIPSE), has documented that although these tests do not measure precisely the same things, when the analysis is conducted on the school level, all of the tests order the schools similarly, regardless of which test or response format is used (Klein, Liu and Sconing, 2009; Shulenburger and Keller, 2009). VSA continues to give institutions a choice among the three tests. To allow time for implementation, VSA gives institutions four years to post results on College Portraits. Work is underway for a Voluntary Framework of Accountability for Community Colleges.

The CLA, CAAP and MAPP tests all take a value-added approach to assessing student learning outcomes using a variety of methodologies. Some of these methodologies can produce unexplained test score variation for an individual institution on the same test across years. For example, until last year the CLA estimated value-added by calculating the difference between the average CLA test scores of a cohort of freshmen and a cohort of seniors and comparing it to the expected difference based on entering SAT or ACT scores. Over time, this methodology

---

Graduate School Admission and Professional Licensure Examinations

Students who are completing undergraduate degrees and continuing their studies in selected graduate programs take the:

- **Graduate Record Examination** (GRE),

- **Medical College Admission Test** (MCAT), or

- **Law School Admission Test** (LSAT).

Examples of professional licensure examinations include:

- the **National Counselor Examination** (NCE) or **Licensed Professional Counselor** (LPC) exam taken by students who complete selected certificates and degrees in mental health or social services;

- the **National Registry of Emergency Medical Technicians** (NREMT) exam taken by emergency medical technicians and paramedics; and

- the **National Council Licensure Examination for Registered Nurses** (NCLEX-RN) exam taken by aspiring nurses.

Other exams exist for pharmacists, respiratory therapists, dental assistants, laboratory technicians, teachers, and many other occupations.

produced wide year-to-year score variability within individual institutions that could not be explained by changes in institutional programs or practices. As of 2009-10, the CLA began using a new methodology that compares actual senior CLA scores to expected senior CLA scores. The expected scores are now being estimated using the school's average SAT or ACT score and the actual average freshman CLA score. They expect this methodology to produce less variability in year-to-year institutional scores. This new methodology has not been in use long enough to assess its efficacy. Therefore, although the appropriate use of the CLA, CAAP and/or MAPP tests can provide important information to support program improvement and student achievement at the institutional level, the Working Group does not think it is appropriate to recommend their required use across the public higher education system in Massachusetts. They appear in the model and plan as direct measures of learning for those campuses that choose to include them among their assessment measures, not as suggestions for system-wide assessment.

Academically Adrift: Limited Learning on College Campuses

The Collegiate Learning Assessment test was recently used as a central component in research conducted by Richard Arum and Josipa Roksa to assess student learning outcomes in critical thinking, complex reasoning and written communication on 24 campuses between 2005 and 2009. Their assessment of student learning during the first two years of college as measured by the CLA test was reported in *Academically Adrift: Limited Learning on College Campuses* (2011). Arum, Roksa and Esther Cho expanded this research to study student learning gains over four years and reported their major findings and recommendations in *Improving Undergraduate Learning: Findings and Recommendations from the Social Science Research Council – Collegiate Learning Assessment Longitudinal Project* (2011). Most significantly, they conclude that 45 percent of students do not demonstrate significant improvements in learning as measured by the CLA during the first two years of college, and 36 percent do not demonstrate significant improvements over four years of college (Arum, Roksa & Cho, 2011, p. 4).

While most of the variability in student learning found by Arum, Roksa and Cho was within individual colleges and universities, the studies documented significantly higher gains in some institutions that have higher expectations and require more reading, writing, and study time. Although these institutional differences are strongly related to admissions selectivity, high- and low-performing students can be found at all institutions. Achievement gaps persist for students whose parents have lower educational attainment levels and continue to grow for African Americans. Higher gains were found for students majoring in traditional liberal arts fields, including social science, humanities, natural science and mathematics, than other majors (Arum, Roksa & Cho, 2011, p. 9-11).

While these findings are clearly concerning for the higher education community and other stakeholders as well, Arum, Roksa and Cho caution that there are no simple solutions and express concern that their conclusions may lead to recommendations to mandate federal or state accountability systems similar to those introduced in K-12 education (Arum, Roksa & Cho, 2011, p. 12-13). On February 3, 2011, Richard Arum addressed the National Advisory Committee on Institutional Quality and Integrity, which advises the U.S. Secretary of Education. At that time he stated:

> "Given the problems of limited learning on college campuses and the recognition that the existing system of accreditation and eligibility for student aid has to date

proven inadequate to the task of addressing these problems, many well-intentioned actors may be tempted to consider imposing an external accountability system on higher education that requires standardized assessments with associated institutional sanctions and rewards attached. Roksa and I believe that such a system would be ill advised at this time. We believe that the existing measures available are not adequate to base an accountability system upon and unintended negative consequences resulting from the introduction of such a system would likely be quite pronounced. However, individuals and institutions must take greater responsibility for addressing the problems of limited learning on college campuses, with accountability best operating through existing governance structures at lower levels of the system" (Arum, 2011).

Arum recommends that the federal (state) government encourage strengthening the existing efforts of institutions to assess student learning outcomes on a broad set of indicators that would include general higher order skills, subject-specific knowledge and other academic competencies; design plans to improve learning outcomes; and monitor progress toward the specific goals set in their improvement plans. He states that "the most useful role for the federal (state) government to play … is to support the advancement of a research infrastructure and the development of appropriate instruments to assess undergraduate learning outcomes and to provide incentives through school improvement grants to encourage colleges and universities to design programs to enhance undergraduate learning and demonstrate measurable improvements in student learning" (Arum, 2011).

It is important to note that, since *Academically Adrift: Limited Learning on College Campuses* and *Improving Undergraduate Learning: Findings and Recommendations from the Social Science Research Council – Collegiate Learning Assessment Longitudinal Project* were just published in January 2011, the higher education and institutional research communities are just beginning to review and comment upon the research. Although the findings are being widely reported, some experts in the field are critical of the research and its conclusions. Alexander Astin, a professor emeritus of higher education and organizational change at the University of California and author of *Assessment for Excellence: the Philosophy and Practice of Assessment and Evaluation in Higher Education* (1991), described serious flaws in the study's methodology and statistical analysis in the February 14, 2011, edition of the *Chronicle of Higher Education.*

The Working Group shares many of the methodological concerns raised by Astin and agrees with Arum that the study results should not cause institutions or state-level leaders and stakeholders to use learning outcomes assessment as a means for apportioning blame or withdrawing financial support. Instead it should propel us to re-energize and re-focus the academy on the core of our mission—undergraduate student learning. We believe that the collaborative effort of public colleges and universities to focus on improvement in student learning through evidence-based and curriculum-based assessment, even as we engage in assessment that allows for state-level comparability and public accountability is the best way to produce the best-educated citizenry for the Commonwealth of Massachusetts. Assessment mandates that yield only compliance behavior will not serve us well for the long term.

The Collegiate Learning Assessment (CLA), the Collegiate Assessment of Academic Proficiency (CAAP) and the Measure of Academic Proficiency and Progress (MAPP) could be helpful instruments for colleges and universities in Massachusetts who wish to use them to examine the performance of their own students and use the results to theorize and test out alternative

explanations and alternative teaching and learning approaches that might lead to program improvement. This is the use of CLA recommended by Richard Shavelson, one of the test designers and a prominent advocate. The Working Group remains concerned, however, about the use of CLA for comparative purposes across states, and we remain convinced that it is not appropriate it use it, or any single test, for high-stakes comparative purposes.

4. Indirect Assessment

Indirect assessment means measuring student learning not by looking "directly" at their work (for example, on tests, papers, or presentations), but rather by using "indirect" means, such as surveys of students, employers, and alumni; focus groups; or exit interviews. While indirect assessment may not specifically and accurately measure individual student mastery of knowledge and skills such as writing, quantitative reasoning, or cultures around the world, it can help describe how effectively individual colleges and universities, and entire state systems, are using best practices in undergraduate education, such as active and collaborative learning in—and out of—the classroom, levels of student-faculty interaction, and support services for students, like academic advising.

Individual colleges and universities employ a wide variety of surveys and focus groups to gather and analyze indirect assessment data that is important to them. Two instruments that are commonly used to indirectly measure student learning outcomes are already in use by most public colleges and universities in Massachusetts: the National Survey of Student Engagement (NSSE) and the Community College Survey of Student Engagement (CCSSE).[2]

The strengths of indirect assessment to meet the challenges of a statewide system for learning outcomes assessment include the easy ***comparability*** of participating states and institutions with one another, often available national benchmarks and the resulting public ***accountability***.

NSSE and CCSSE in particular, have several strengths that make them especially useful to campus programs and system-wide learning outcomes assessment. They share the following important features and characteristics:

- The soundness of their academic underpinnings and sophistication of their instruments. Both are based on voluminous research on the impact of college on students and student learning and from empirically vetted, high-impact practices in higher education. (Pascarella & Terenzini, 2005; Carini, Kuh & Klein, 2006; Kuh, 2008)

- The ability to provide diagnostic information to support improvement in undergraduate education. Items on both surveys ask students to report on their engagement: such things as how much time they spend preparing for class; how much reading and writing they do in connection with their classes; the extent of their interaction with faculty outside of class; their participation in internships, community service and study abroad; and the extent to which they feel the institution helps them succeed academically and socially (Kuh, (2008). Many instrument items are common to both NSSE and CSSE, while others are designed

---

[2] Information about NSSE and CCSSE taken from the websites of both organizations: http://nsse.iub.edu/ and www.ccsse.org.

for the differing characteristics and academic environments of four-year and two-year institutions.

- The built-in benchmarking that allows for longitudinal analysis and comparisons with other institutions. Benchmarking builds a culture of evidence and allows a focus on performance and success. NSSE and CCSSE each use five benchmark categories with the first three held in common:

NSSE	CCSSE
Level of Academic Challenge	
Active and Collaborative Learning	
Student-Faculty Interaction	
Enriching Educational Experiences	Student Effort
Supportive Campus Environment	Support for Learners

The differences in the last two categories reflect the fact that most four-year institutions feature full-time study in residential environments while community colleges usually serve commuter students, a great many of whom study part time.

- Both NSSE and CCSSE encourage the voluntary formation of consortia which enable within-consortia comparisons and specially developed survey items. They also allow for peer institutional and cross-state comparisons. Fourteen states, including Massachusetts, are listed on the CCSSE website has having statewide or state-based consortia for 2011. NESSE has many consortia made up of special-purpose or special-identity institutions (e.g., Catholic colleges, public liberal arts colleges). It also has a number of statewide or state system-based consortia (e.g., Pennsylvania State System of Higher Education, Connecticut State Universities, South Dakota Publics).

- Both NSSE and CCSSE are committed to principles of public accountability and the sharing of information. Colleges know in advance that CCSSE results and benchmark scores will be publicly presented on the CCSSE website. NSSE results and benchmark data are sent to participating campuses but made publicly available only by institutional authorization. Both NSSE and CCSSE have strong national organizations and widely used instruments.

NSSE was created in 1999 when a group of higher education leaders with support from the Pew Charitable Trusts initiated a project to measure the quality of undergraduate education that would be better than ranking systems based on faculty reputation and/or student selectivity. It is now a large organization with a prestigious National Advisory Board, several technical committees, and extensive foundation support. It is directed by Alexander McCormick and headquartered at Indiana University in the Center for Postsecondary Research and Planning. It remains closely identified with George Kuh, founding director of the Center for Postsecondary Research and NSSE at Indiana University.

The NSSE survey is now the most widely used annual survey of undergraduates in the country. The 2010 survey was administered to 363,859 students attending 603 colleges and universities, and the cumulative total of students taking NSSE since 2000 is 2,321,085 for 1,452 participating

colleges and universities. In Massachusetts, four state universities (Bridgewater State, Mass College of Art and Design, Salem State and Worcester State) and the four undergraduate campuses of University of Massachusetts used NSSE in 2010. Eighteen private colleges used NSSE as well. Cumulative totals would add two more state universities (Fitchburg and Framingham) and a large number of additional private colleges to the list.

NSSE produces annual reports based on findings for all participating institutions which are intended to highlight changes in student engagement over time. Other services include specialized reports (e.g., the psychometric properties of the survey, best student engagement practices, program improvements made by campuses as a result of NSSE), research projects, and workshops for survey users.

CCSSE was established in 2001 as a project of the Community College Leadership Program at the University of Texas at Austin. It is now part of the Center for Community College Student Engagement but still closely affiliated the University of Texas at Austin. It is directed by Kay McClenney, a leading authority on community college education. The development and continuation of CCSSE has been supported by many national foundations, and it has an impressive National Board with some planned overlap with the NSSE board.

CCSSE was intentionally built on the base established by NSSE, and both organizations continue to work in close partnership. The CCSSE survey instrument and benchmarks draw on the research base that supports NSSE but also on the research on community college student learning, learner-centered and learning organizations, effective educational practices in community colleges, and the needs and special characteristics of community and technical colleges and their students. The Center publishes national reports, conducts ongoing research, and offers a variety of resources and services.

> **IN THE WORDS OF THE FACULTY**
>
> *"We need to be thoughtful about how we look at data, better data is often more complex."*

CCSSE is nearly as widely used as NSSE. The 2010 survey was administered to 400,000 students on 658 campuses. Since the first survey in 2002, CCSSE has surveyed more than a million students in over a thousand institutions. Community colleges in Massachusetts have participated in CCSSE in various numbers over the years and formed a consortium for 2007 as well as for 2011.

There is a growing body of research on the relationship between NSSE/CCSSE measures of student engagement and measures of student learning and student success. Here are three examples:

- A recent analysis of data from the Wabash Study (Blaich, 2007), a longitudinal investigation using a variety of direct measures of the institutional experiences that enhance growth in educational outcomes, found positive associations between NSSE benchmark scores and results of direct measures of seven liberal arts learning outcomes at the end of the first year of college, independent of the differences across participating institutions (Pascarella, Seifert & Blaich, 2010; Blaich & Wise, 2011). Although the Wabash Study is small, it is important because it suggests that selected NSSE results can serve as proxy measures for growth in important learning outcomes.

- A CCSSE validation study examined the relationship between CCSSE engagement measures and "easily verifiable" outcomes such as course completion, GPA, persistence

and completion measures (McLenney, Marti & Adkins, 2006). Finding positive relationships among the measures, the researchers concluded that CCSSE serves as "a valuable proxy for student success" (p. 2).

- AAC&U's LEAP project examined the results of their Compass I project, a collaborative of three state systems—The California State University, the Oregon University System, and the University of Wisconsin System—and selected campuses within these three systems that together worked to re-map general education using new approaches to program design and learning outcomes assessment. The approaches were drawn from LEAP Principles of Excellence and Essential Learning Outcomes, from Kuh's work with high-impact educational practices and NSSE, and from lessons learned in AAC&U's "Making Excellence Inclusive" initiative. "Making Excellence Inclusive" concentrates attention on the success of students from underserved groups. Studies of Compass I and reviews of other research show that high impact practices are related to higher grades, higher persistence rates, intellectual gains, greater civic engagement and increased tolerance for, and engagement with, diversity and increased interaction with faculty and peers (Association of American Colleges and Universities, 2011; Brownell & Swaner, 2010; Finley, 2009). The pattern was especially clear when two or more high-impact practices were involved. From this work, AAC&U notes that NSSE data "is serving reliably across three systems as a proxy for learning and student success" (Finley 2009 via personal communication with Susan Albertine, Association of American Colleges and Universities).

Although NSSE and CCSSE are measures of student engagement and only indirect measures of student learning, their widespread use throughout the public system, the advantages of benchmarking and consortia, and close relationships to learning outcomes make them important tools in a multiple-measures approach to learning outcomes assessment. They bring together faculty and staff, the curriculum and the cocurriculum and well serve the assessment-for-improvement and the assessment-for-accountability needs of both campuses and DHE.

5. Composite Learning Indicator

The three elements of the "triangulation" approach to statewide learning outcomes assessment (indirect, direct, and embedded assessments as shown in Figure 1) each produce results that may

be used, in varied ways, to improve learning at the classroom and campus levels. While **improving learning** is the most vital function of outcomes assessment, communicating the results of teaching and learning to the public (the **accountability** feature) is also important.

In Chapter One, we recommended the use of a dashboard for reporting the results of learning outcomes assessments. Campus dashboards would allow for results presentation consistent with the pyramid structure of the model. Dashboards can also be used for reporting state-wide results by segment for community colleges, state universities and the University of Massachusetts undergraduate campuses, or for public undergraduate campuses taken as a whole. The Working Group suggests that dashboards would provide the best means of outcomes results presentation.

<table>
<tr><td>

IN THE WORDS OF THE FACULTY

"Much is already working well in assessment."

</td></tr>
</table>

The Working Group is aware however that a desire for a single statewide measure of student learning for the Vision Project was an important element in the Commissioner's charge to the Working Group. We continue to recommend against any single measure or test for the Commonwealth. As an alternative, however, and as a way of communicating simply and clearly the combined accomplishments of Massachusetts' public colleges and universities on multiple measures of learning, a **Composite Learning Indicator** could be developed and used with the Vision Project. The composite learning indicator would be constructed from the three types of assessments available in the pyramid in our model. The Working Group is aware that it will require substantial effort on public campuses to develop and implement assessment measures that support their own assessment programs directed toward learning improvement **and** that can be used in the composite learning indicator, but we believe that with sufficient time and resources and with effective connection to LEAP learning outcomes and rubrics and to NSSE/CCSSE, it will be possible to accomplish this. Further elaboration of the composite learning indicator is needed as part of plan implementation.

The Working Group also explored the possibility of a composite indicator of educational capital which would combine learning, educational attainment and success, and post-graduation employment measures to provide a broader picture of the educational contributions that public higher education provides to the Commonwealth. Because it goes beyond the specific area of our charge and overlaps with other outcome areas of the Vision Project, we have not included this composite indicator in our Plan for Statewide Learning Outcomes Assessment. But we have included it in Appendix D, together with a description of its genesis in work intended to describe the educational capital produced in all 50 states. The concept may prove useful to the Vision Project.

6. State Partnerships

Throughout this report, we have noted the difficulties associated with comparing the results of learning outcome assessments in Massachusetts with those in other states. We have noted many methodological and substantive issues related to comparability and found wide variation in assessment practices among the states that would add to the difficulty. It is also the case that of the 12 LEAP learning outcomes, only four can be assessed using the CLA, MAPP, or CAAP tests: critical thinking, analytical reasoning, problem solving and written communication, and we have argued that none of these tests would be suitable for statewide use or state-to-state comparisons. Given this, the Working Group has been discussing other ways to achieve the

Vision Project goal of establishing Massachusetts' standing with respect to other states in the area of student learning.

One possibility is the creation of state partnerships. Massachusetts would develop relationships with a number of other states and work with them to develop appropriate and meaningful comparisons that could be reported publicly. It would be important that partner states be similar enough to Massachusetts in the size and structure of their public higher education systems to make the comparisons meaningful and that partner states share our commitment to learning outcomes assessment that can be used both for campus improvement and for accountability. Ideally, state partners would be chosen from states which also have made commitments to LEAP learning outcomes and rubrics as part of their assessment planning.

In early March four members of the Working Group and a faculty member from Salem State University attended a LEAP States Summit sponsored by the Association of American Colleges and Universities. Fourteen states sent teams to the summit including all six current LEAP States. We shared our ideas for state partnerships with representatives from other states and several expressed interest and willingness to follow up with partnership explorations. It will take time to develop appropriate state partnerships and forge agreements about comparable metrics and it will take time for implementation as well, but over the long term, the Working Group believes that this approach holds the best promise for meaningful comparisons across states.

> **IN THE WORDS OF THE FACULTY**
>
> *"In communicating with the public at large, we must overcome the conflation of 'testing' and 'assessment.' In the worst case scenario, standardized testing is equated with assessment, because it leads people to assume that it is the only valid way to evaluate students' learning. This is, of course, a false assumption. Standardized testing is a small subset of testing, which itself is a small subset of assessment. Since I tend to think visually, here is an illustration:"*
>
>
> All assessment
> Testing
> Standardized testing
>
> —*Charlotte Gifford, Interim Dean of Humanities, Greenfield Community College*

AAC&U has several new LEAP-related initiatives that will be helpful to our efforts in state partnership. A Collaborative on Authentic Assessment of Learning (CAAL) will provide an assessment data repository, a virtual community where institutions can upload, share, and discuss their campus assessment results using the VALUE rubrics. The collaboration among e-portfolio vendors and campuses will expand communication across campuses regarding common instruments, compelling findings, potential benchmarks for student success, and best practices of assessment using rubrics and e-portfolios. Discussions with Terrel Rhodes in Chicago sparked interest in the possibility for state system participation in CAAL. Advantages for campuses and the system should continue to be explored as CAAL develops.

Other new initiatives include a VALUE Rubric Reliability Study in which faculty volunteers across selected disciplinary areas are scoring student work samples and developing reliability scores to

assess the degree of shared understanding of rubrics across and within disciplinary areas; a project to develop online tools for demonstrating how to sample student e-portfolio work; and work to develop criteria for selecting appropriate rubrics and effective strategies for communicating results.

The Working Group recognizes that our plan for statewide learning outcomes assessment is a very ambitious plan but we believe that it is the best plan for "doing learning outcomes assessment right." By building on the work of faculty and staff on all campuses, on campus-based approaches to learning outcomes assessment and on strong collaborations among campuses and with the Department of Higher Education we can produce results that can be used for improvement and accountability. We believe that the Commonwealth of Massachusetts can achieve real leadership by such an approach.

Chapter Three
Resources and Implementation

This chapter considers resources and implementation associated with the Plan for Statewide Learning Outcomes Assessment. It recognizes that effort and financial resources will be required at academic program, campus and state levels to strengthen existing work in learning outcomes assessment, build cultures of evidence, and link results to program improvements that foster student learning. We suggest a number of ways that campuses can work together with DHE to foster collaboration and commitment that will help realize the ambitious Vision Project goals in the area of student learning and learning outcomes assessment. This chapter gathers suggestions from throughout this report, and from our Phase One report, and presents them for consideration as implementation steps. While it does not treat the subjects of resources and implementation in detail, the intent of this chapter is to recommend a direction for future activity by identifying key issues and concerns and making suggestions for moving forward.

Resource Implications

Context: Costs and Benefits, Purposes and Priorities

The statewide model for learning outcomes assessment relies heavily on outcomes and assessment work on each of the 28 undergraduate campuses in the public higher education system and on their united effort, in collaboration with the Department of Higher Education, to use results for program improvement, track progress for the Vision Project, and respond to accountability concerns. This approach to assessment is far more labor intensive than the alternative approach of a single statewide test. But, for reasons explained elsewhere, the Working Group believes it is a far better approach for making genuine improvements in student learning.

Because it is more labor intensive and involves multiple assessment measures, however, the approach is also more costly. In an era of serious fiscal constraint, increased costs can only be justified by genuine benefits (Wellman, 2010). The Working Group took seriously the goals of the Vision Project that "Massachusetts produce nationally leading educational results" and takes seriously our charge to "identify ways to add strength and capacity to campus and system-level assessment efforts as one means of improving the quality of student learning in public colleges and universities throughout the Commonwealth" (Phase One Report, p. 2). Our plan reflects our conviction that public higher education can best achieve the ambitious outcomes called for in the Vision Project and best serve the needs of the Commonwealth by building from the broad base of faculty and staff on campuses, by focusing on evidence of learning, by using assessment results for improvements in teaching and learning, and by informing the public about our progress. The Working Group is convinced that there are genuine, long-term benefits for students, for institutions, for the public higher education system, and for the Commonwealth from this approach and that these benefits do justify the increased levels of public appropriations necessary to make effective learning outcomes assessment possible throughout the system. Stated simply, in the plan we have

> **IN THE WORDS OF THE FACULTY**
>
> *"If assessment becomes top-down, the enthusiasm will decrease."*

proposed, learning benefits accrue to students, institutions and the Commonwealth and make national leadership for Massachusetts possible.

Because the resource needs will be extensive and resources are scarce it is also important to be clear about purposes and priorities. The Working Group suggests that campuses and DHE focus on improving teaching and learning by ensuring that there are excellent, evidence-based learning outcomes assessment programs on each public campus, by creating a strong system-wide network of assessment experts to facilitate shared learning and best practices and by developing and using assessment measures that allow for meaningful institutional, segmental and system-wide comparisons, transparency and accountability.

Although they will be costly because they will require additional efforts on most campuses, embedded, direct and indirect assessments all contribute to improved student learning by making it possible for campuses to identify areas for improvement and to make changes in curriculum, programs or support practices. They also contribute to student learning indirectly because they allow for comparisons for institutions and the Commonwealth. Institutions can track their progress with respect to their peers, and DHE can compare overall Massachusetts progress with that of other states. In both cases accountability obligations are recognized and discharged, and the results can be used to identify specific areas for program improvement.

Estimating Resource Needs.

The plan we have proposed will require new or reallocated assessment resources for campuses who intend to improve their capacity for evidence-based assessment through continued refinement of learning outcomes, curricular alignment activities, use of multiple and embedded assessment measures, and increased attention to feedback loops for curricular and teaching improvement. There are also resource implications for campuses and DHE associated with incorporating additional mechanisms for system-wide collaboration and for the development and analysis of measures that will be used for peer institutional comparisons and state comparisons.

While detailed cost estimation work is beyond the scope of this report, we can suggest some basic strategies and models, drawn primarily from the work of Swing and Coogan (2010) for NILOA, that could inform this work on campuses and at the DHE. We do believe that recognizing resource implications before going forward is important for several reasons. If resource issues are not addressed in advance, this plan could have the effect of creating large, unfunded mandates for campuses and the DHE and/or lead to uncoordinated "start and stop" initiatives that will waste resources over the long term and not produce the levels of quality in student learning toward which we aspire as a system. Also, a failure to be clear about resources for this endeavor could create a situation in which some campuses do less expensive assessment to save money while others do better assessment because they can afford it. This will not lead to the position of leadership in student learning and learning outcomes assessment to which the public system aspires.

Examples of potential estimation strategies include:

- Campuses may wish to utilize project management models that estimate all additional costs associated with making desired additions or modifications to current assessment programs and practices and then set priorities and timelines that ensure that resources will

be available to complete the work. This allows priority work to move forward—but at a realistic pace.

- For estimations associated with planning for the addition of new assessment tests and surveys, Swing and Coogan (2010) examined numerous institutions and identified several direct and indirect costs that need to be considered. Their model includes:

  - o surveys and tests
  - o software and hardware for assessment tracking, scoring, and analysis
  - o postage
  - o student incentive awards
  - o third-party scoring and/or reporting services
  - o access to benchmarking/peer comparison data (for example, the National Student Clearinghouse) (Swing & Coogan, 2010, p. 8).

- A more inclusive model for estimating costs for adding new assessment measures, developed and tested in different types of institutions by Peter Ewell and Dennis Jones (1986) at NCHEMS, builds in the associated personnel and institutional costs with these categories:

Institution Type	Instrument	Admin. of Survey	Salaries and Benefits	Overhead	Total

Source: Ewell & Jones, 1986, as cited by Swing & Coogan, 2010, p. 10.

In the model, "Instrument" refers to the costs to develop and/or purchase survey instruments (e.g., CCSSE, NSSE) and to score the surveys. The "Administration of Survey" item includes costs to announce, administer, and mail the surveys. "Salaries and Benefits" includes costs for salaries and benefits of testing/measurement specialists, secretaries, and student workers, etc. "Overhead" includes costs for consulting visits (fee and travel), data analysis, reports, and office expenses.

> **IN THE WORDS OF THE FACULTY**
>
> *"Better assessment will require a mix of external help and internal resources."*

The Ewell/Jones model, or other similar models, can be used for estimating costs associated with adding CCSSE or NSSE to campus assessment programs or for adding other new elements at campus and system levels. Using the same estimation models across the system would provide a clear and distinct process that can be uniformly applied and used to estimate costs for the system as a whole. Outside of the personnel costs the expenses can be determined with confident accuracy for initial startup. Personnel costs will be estimated in the startup but could be more accurately assessed as an institution gains experience administering the assessment.

The model could be useful for estimating embedded assessment costs at the campus level as well but in this case different campus approaches to assessment and different technologies challenge system-wide uniformity and make estimation more complicated.

Faculty are building their outcomes, (e)portfolios, and assessments on various technology platforms that could be difficult to coordinate within campuses and across the system.

- Another approach emphasizes personnel cost estimation. Faculty, academic administrator, and staff time can be approximated and costs estimated from salary data when these individuals are involved in planning, development, management, oversight and/or administration of learning outcomes assessment measures and activities within or beyond their regular responsibilities. One model proposed that:

    o Faculty time needs range from 10%, if the faculty role is limited to consulting and advising; to 30%, if new/specific assessment instruments are developed by the campus; to 50%, for assessments requiring faculty evaluation of live performances, portfolio reviews, and other labor intensive methods.

    o About 30% of an assessment administrator's time is needed to conduct a campus-wide assessment of a learning outcome goal.

    o Estimates for other administrators and staff can be based on the percentage of their yearly time that is devoted to assessment (Harper 2009, as cited by Swing & Coogan, 2010, p. 8).

Additional Considerations

The Working Group discussed many other issues that have important resource implications. We believe it is important to consider the following:

- This plan relies heavily on faculty work and the work of professional assessment administrators and institutional researchers. The Working Group believes that extensive faculty involvement and a professional assessment staff is the correct strategy for achieving lasting, positive impact on student learning. In our meetings with faculty we were reminded that community college faculty members are expected to teach five courses per semester. Additionally, many institutions now use adjunct/contingent faculty or graduate students for a large proportion of their undergraduate teaching, and these individuals are paid on a part time and often per-course basis. This has significant implications for the assessment of student learning. For each segment (i.e., community college, state university, University of Massachusetts) and for the system as a whole, overall faculty size, faculty teaching loads and collective bargaining contracts will need to be made consistent with expectations in the area of learning outcomes assessment.

- More assessment activity on campuses and new components and analyses will increase workloads and create needs for additional professional administrators in assessment and institutional research. In our Phase One report we noted that "Massachusetts college presidents report both their strong support for student learning outcomes assessment and their concern about funding to support its implementation. Only 36% of our community colleges, 33% of state universities and 50% of the University of Massachusetts campuses have a Director of Assessment (Phase One report, p. 34).

- Another set of activities that will incur a cost are those associated with professional development for faculty and staff. Professional development activities are often provided by external consultants who offer workshops and planning assistance. New networks for

cross-campus collaboration and sharing of best practices should help to reduce overall costs in this area, but professional development resources for campuses will still be needed.

- We recognize that in addition to the learning assessment elements in this plan, campuses and DHE conduct other assessments specific to their programs and for many other aspects of institutional effectiveness. How campuses and DHE balance other assessment activities and obligations with those in student learning is an important issue that should be addressed.

- Consideration should be given to building campus assessment on the same technology platforms across the system. This could improve coordination, increase transparency, and provide cost savings.

- Collaborative approaches to planning for assessment offer opportunities for sharing best-practice information and for experimental and pilot studies prior to full-scale implementation. These also create possibilities for cost savings for institutions and the system over the long term.

Availability of financial resources will heavily influence the overall quality of our assessment programs and our ability to move forward.

Implementation Plan

The intent of WGSLOA Phase II recommendations is to achieve a system of assessment, statewide, that promotes continuous improvement and accountability in public higher education.

IN THE WORDS OF THE FACULTY

"Use data to inform faculty professional development and not for punitive measures."

It is our aim to become a national model whereby our institutions of higher education will share best practices and develop common models of assessment while maintaining the campus autonomy that allows for the rich diversity of programs we enjoy across our campuses.

Extensive work is already underway at the campus level in response to expectations of accrediting bodies and the interest of faculty and staff to ensure quality education for our students. The DHE also has gathered performance measures annually for many years. These metrics provide a foundation for the measures that can be used in the future to determine the effectiveness of our programs and to compare our work to that of public higher education in other states.

There are four major considerations that we have discussed and that deserve repeating as we outline our plan for implementation:

- The first consideration is that our academic programs are the work of our faculty, and their full participation in the assessment of programs is essential to develop strong models of evaluating student learning outcomes.

- Second, as noted in the resources section, the work of assessment requires financial resources at both the campus level and statewide. In-depth assessment is comprised of multiple measures of student learning, much of which demands a qualitative approach and therefore adds to the work load for our faculty and professional staff. While assessment is

a component of the work of faculty and designated professionals, building a statewide model cannot be accomplished by placing additional expectations on personnel whose current work responsibilities are already extensive. Our university and community college faculty have heavy teaching responsibilities along with advising and committee responsibilities. Additional resources for faculty and for assessment personnel will be needed.

- The success of this initiative depends on faculty ability to implement the assessment models. It is essential that they have the opportunity to gain knowledge in assessment to achieve the goals. Professional development is a critical aspect of this plan, and resources for faculty professional development will be needed.

- Finally, it will be important to knit the new assessment efforts with existing accountabilities to accrediting agencies both at the institutional level (NEASC) and discipline level (specialized program accreditations such as ABET, NCATE, and others). With resources limited, we cannot duplicate our efforts with multiple reporting requirements. It is essential that we are both efficient and effective in assessing the quality of our programs.

The implementation plan is contingent on the active participation of all of undergraduate public higher education institutions and their willingness to join in collaborative efforts with the DHE. The plan also requires DHE to recognize the work that is already underway at the campus level and work to allow for the uniqueness of each campus to be incorporated into the final evaluation of campus work.

Timeline and Action Steps

This implementation plan is a five-year plan. Where possible, specific timelines are identified; however, the timing of various components will need to be determined as the work gets underway.

I. **LEAP - National**

    A. Become a LEAP State (AY12) – it is the most attractive path to a form of comparison across states and the LEAP Essential Learning Outcomes are currently reflected in most institutional outcomes and provide a framework for the Statewide Plan.

    B. Establish Massachusetts LEAP Team. (AY12)

        1. Include representation from each segment (University of Massachusetts, state universities, and community colleges) and DHE.

        2. Have members of MA LEAP Team also be liaisons to the groups identified in the following sections.

    C. Begin the work defined by the LEAP national leaders to participate in the national effort and craft Massachusetts' particular approach to using LEAP as a framework. (AY12)

II. **Resources for *A Plan for Statewide Learning Outcomes Assessment***

    A. Develop a detailed resource plan (in the model of project management planning) that outlines direct and indirect resources needed by the DHE and campuses to achieve the goals of the Vision Project.

1. Consider new costs of assessments such as all campuses administering the NSSE or CCSSE.

2. Determine method for quantifying the additional time needed from faculty, assessment professionals, DHE and others to further develop the statewide plan and assign a monetary value to the personnel costs.

3. Develop an operating budget for meetings, travel, educational resources, etc.

B. As part of the resource plan, evaluate the proposed time lines for objective attainment based on the resources available.

III. Campus Engagement

A. Engage Provosts/AVPs[3] from each of the 28 campuses (or their designees) to lead work to:

1. Develop plans for engaging faculty and faculty governance in decision-making regarding assessment (and campus buy-in). (AY12)

   a. For each segment (and therefore contract) identify any issues of contractual concerns; workload; etc.

   b. Develop a process to achieve a blend of leadership between professional assessment personnel and faculty to guide efforts on campus and statewide (governance committee and other activities).

2. Have campuses analyze their local student learning outcomes as compared to LEAP outcomes to determine areas of similarities and differences. (AY12)

3. Determine applicability of LEAP as the shared framework to assess outcomes and/or establish common, though not necessarily identical, outcomes for public higher education in Massachusetts. (AY12)

4. Endorse LEAP's commitment to liberal education and its essential learning outcomes as required by the LEAP State Guidelines. (AY13)

5. Implement the Plan for Statewide Learning Outcomes Assessment that uses multiple measures of student learning including indirect, direct, and embedded assessment (hopefully with LEAP Rubrics). (AYxx)

6. Inventory resource needs on campuses to achieve a model of assessment as presented in the goals of the Vision Project. (AY12)

B. Establish a network of assessment leaders from each of the 28 campuses and with representation from DHE.

1. Invite participation from all campuses to have a designated assessment leader for that campus. (AY12)

2. Appoint a faculty chair, by the Commissioner, upon the advice of an advisory body of campus presidents/chancellors representing the three segments. (AY12)

---

[3] The work of assessment rests with the faculty and professional assessment staff. The provosts/AVPs are the individuals on each campus that can lead this effort and insure that the work of the statewide effort is integrated with the work on the campus level. Sarah—isn't it "who" can lead this effort—you decide.

3. Work plan

    a. Define current state of assessment on campuses (AY12)

        i. Develop a baseline body of evidence of:

- Current best practices across all campuses

- Common areas of challenges (weaknesses)

- Use of surveys such as NSSE/CCSSE

        ii. Identify current mechanisms of improving assessment on campus

- Local efforts such as having an identified assessment leader on campus working with faculty

- Professional networks/conferences to share information and learn from each other

- Tapping regional/national leaders in assessment for individual institutional review/guidance

    b. Develop professional development plan to advance knowledge and practices on campuses (AY12)

> **IN THE WORDS OF THE FACULTY**
>
> *"Partner with other similar states to improve assessment for all, instead of striving alone to be 'Number 1.'"*

        i. Offer workshops/conferences to share information

        ii. Develop discipline specific teams to offer models in their areas of education

        iii. Provide access to information offered regionally/nationally

    c. Along with Institutional Research colleagues,

        i. Provide an analysis of how data from CCSSE/NSSE could be used as one measure of student achievement and the value/practicality of establishing consortia within segments (AY12)

        ii. Propose student learning outcome measures for a Massachusetts dashboard and composite learning indicator. These measures should allow each campus to also be compared to their institutional peer group. (AY12 – draft to test)

        iii. Develop plan/model/process for selecting peer institutions and peer states (AY12)

        iv. Having received recommendations from the LEAP group, refine what information needs to be collected by DHE to provide evidence of achievement towards Vision Project goals within LEAP framework and with particular consideration of existing reporting requirements to NEASC (AY13)

    d. Periodically (quarterly) present information to the Advisory Body of Presidents/Chancellors and Commissioner for the statewide effort and to

Provosts/AVPs and Presidents for campus activities for review, feedback, and direction.

The Working Group recognizes that the premise of the Vision Project is to set clearly ambitious outcomes and to measure our progress toward these outcomes on an annual basis. While we understand and appreciate this need, we want to emphasize that it creates many difficulties in the area of learning outcomes assessment. Achieving national leadership in the quality of student learning and in the assessment of student learning cannot be accomplished in short order. We ask the Commissioner to recognize that changes to improve learning outcomes assessment programs and improve student learning will take time: to establish learning outcomes and/or make LEAP-related adaptations; to align outcomes with curriculum; to modify or develop assessment measures; to collect evidence and analyze results; to use results to develop plans for program improvements; to implement the plans, evaluate the results and plan again for improvements; to add student learning measures to peer-institution comparison frameworks; to add measures for the Composite Learning Indicator. All of this is consistent with the Culture of Evidence approach recommended in our Phase One report and requires a long natural cycle of evidence-based assessment and improvement. Our suggestion for the timing dilemmas is to develop clear priorities for activities associated with the statewide plan and build the priorities into the Vision Project as action steps over time as resources permit.

Appendix A

Executive Summary from Phase One Report from the Working Group on Student Learning Outcomes and Assessment

The Working Group on Student Learning Outcomes and Assessment was established by Commissioner Freeland in January, 2010 to bring together public campus presidents, provosts, faculty members, students, institutional research/assessment directors and a member of the Board of Higher Education in a collaborative deliberation about one of the components of the Vision Project—student learning outcomes and assessment. The Vision Project calls for Massachusetts to "produce the best-educated citizenry and workforce in the nation" and establishes student learning achievement as one of the seven outcome areas in which the Commonwealth will strive for leadership. The Commissioner asked the Working Group to examine national best practices and current status on Massachusetts public higher education campuses in two phases of work—the first to focus on campus programs for learning outcomes assessment and the second to focus at the state level on issues of public accountability, comparability and transparency.

The Working Group has met seven times since January, 2010 and has completed Phase One of its work. As requested by the Commissioner we have:

- reviewed national literature, research and best practice on campus-based learning outcomes and assessment,

- reviewed the results of a DHE Survey of campus programs sent to public campus provosts in December, 2009 and designed for comparability with a national survey prepared by the National Institute on Student Learning Outcomes and Assessment,

- considered the report of the Commissioner's Advisory Group on Undergraduate Education (CAGUE),

- considered the Liberal Education and America's Promise (LEAP) initiative of the Association of American Colleges and Universities, and

- considered approaches to assessment that value our system of diverse institutions and that contribute to reductions in performance gaps over time among different ethnic/racial, economic and gender groups.

We have met with Commissioner Freeland and submitted a Phase One report. This executive summary covers some of our conclusions and recommendations in abbreviated form. The full Phase One report and appendices will be available on the DHE website by September 2010. A Phase II Report will be available at the conclusion of our work in December 2010.

Major Findings

- Momentum toward greater public disclosure and accountability by colleges and universities and state higher education systems for student learning has been building for several years.

- There is a great deal of consistency between learning outcomes and assessment programs in Massachusetts and patterns throughout the country.

- Most public colleges and universities in Massachusetts have in place a common set of learning outcomes that apply to all undergraduate students, as well as learning outcomes for specific programs. Most Massachusetts public colleges and universities include Communication, Quantitative Reasoning and Critical Thinking among their institutional-level learning outcomes.

- There are substantial similarities among the learning outcomes in place at Massachusetts public colleges and universities, New England Association of Schools and Colleges (NEASC) accreditation standards, CAGUE recommendations, and the LEAP "Essential Learning Outcomes."

- Massachusetts public institutions of higher education have developed student learning outcomes assessment programs that include many of the key components of exemplary assessment programs identified in the literature (Kuh & Ikenberry, 2009).

- NEASC standards, guidelines, and data requirements have exerted a strong influence on the approaches that Massachusetts public colleges and universities have taken to program and institutional learning outcomes assessment. Institutional assessment of student learning outcomes is typically done either for institutional program review or for self-studies for accreditation.

- Good assessment of student learning outcomes in Massachusetts derives from the faculty. The level of faculty involvement in assessment is a key indicator of a successful assessment program. Faculty must be responsible for the development of learning outcomes and involved in designing, scoring, analyzing, reporting, and using assessment data with strong support from administrators and an assessment office. Faculty time, effort, and commitment needed for assessment programs to be successful must be recognized and faculty responsibilities must be adjusted and rewarded accordingly.

- It is entirely unlikely that the outcomes measured by Collegiate Assessment of Academic Proficiency (CAAP), the Measure of Academic Proficiency and Progress (MAPP), or the Collegiate Learning Assessment (CLA) will correlate well with intended learning outcomes articulated locally by the faculty of a given college or university campus. That is, these tests are unlikely to have content validity. If standardized tests are adopted as a common measure across institutions they will supplant the richer, more robust, more evolved, and more useful local assessments that are currently being employed.

- The lack of a sufficiently developed culture of assessment is cited by many public higher education institutions' provosts as a weakness. Assessment is not a one-time event; it must be seen and supported as an ongoing, iterative, cascading process if it is to result in the desired program improvements and better student learning.

- There is a need for more leadership, better coordinated efforts, and access to top notch assessment experts to support the development of top tier assessment programs on Massachusetts campuses.

Recommendations

The Working Group on Student Learning Outcomes and Assessment understands both the desire to create a more centralized and systemic approach to student outcomes and assessment as well as the need to maintain decentralized and campus-centered control over student learning. Within the nuances and complexities of that reality, our Working Group:

1. Sees potential value in aligning our state system of higher education with the LEAP framework, including the ability to benefit from assessment frameworks and practices developed by other LEAP states and institutions, the possibility of applying the LEAP framework to increase the achievements of underserved students, and the potential for making comparisons with other LEAP states in the future.

   We recognize that the identification of learning outcomes is a responsibility of the faculty and recommend that the identification, implementation and assessment of field specific learning outcomes should remain the responsibility of faculty at individual colleges and universities. We recommend that Massachusetts consider becoming a "LEAP State," but only after having consulted with the faculty of Massachusetts public colleges and universities regarding their opinions of the same.

2. Recommends that individual campuses continue their efforts to develop assessment programs and strategies that clearly reflect their mission, academic programs and co-curricular offerings. Planning at all levels of each institution is necessary and critical in determining the student learning outcomes expected and optimal strategies for assessment. Assessment should occur at the course, program and campus levels to maximize data collection and continued improvement.

3. Recommends that each campus review its capacity to provide the leadership, expertise, assessment instruments and support to faculty and staff engaged in campus-based assessment activities and to achieve system-wide assessment goals. This capacity should be improved, if necessary, as funds become available. Campuses are also encouraged to consider taking a regionalized approach to assessment and/or the requisite professional development for faculty and staff as a way to reduce program costs and to increase the accessibility of assessment expertise and opportunities for faculty collaboration.

4. Is strong in its opposition to any high stakes standardized assessment instrument for all students in the Commonwealth's higher education system. It is imperative that any approach to the assessment of student learning be sensitive to and understanding of students, communities and missions of the three sectors of public higher education in Massachusetts.

5. Values the concepts presented in "Creating a Culture of Evidence, An Evidence-Centered Approach to Accountability for Student Learning Outcomes" (Millett, Payne, Dwyer, Stickler & Alexiou, 2009) as a holistic, cost-efficient process to create and maintain the culture of evidence necessary to improve student learning outcomes on campus. The Working Group encourages each public higher education campus to review the model, or similar models, and consider its implementation on their campus.

6. Recommends that the DHE identify sources and create funds to support both experimental approaches and assessment initiatives at the system and campus level that document the link between assessment results and program improvements.

7. Recommends that the colleges/universities, in collaboration with the DHE, engage a high-profile assessment consultant (guru) to provide vision, expertise, and leadership to system-wide student learning outcome assessment and that each campus identify a point-person for assessment to work with the consultant and other institutional representatives. The consultants work would include individual campus consultations on assessment, tailored to the campus needs, and convening meetings of campus-based assessment staff and representatives of DHE for sharing best practices and challenges and developing common assessment data that can be used to provide evidence of the achievements of public higher education across the Commonwealth. The consultant's work would extend one to two years, with ongoing work the responsibility of campus representatives and DHE staff.

8. Recommends that the repertoire of campus assessment approaches include the capability of analyzing the achievement of comparable learning outcomes among different ethnic/racial, economic and gender groups and encourages campus efforts to narrow performance gaps. DHE should play a helpful role in support of this objective by providing consultative assistance as well as templates and tools for ways to disaggregate and analyze data so that performance gap analysis becomes a part of ongoing institutional research activity on each campus. Given the high costs involved in achieving sufficient sample size to allow for disaggregated analysis, system-wide and segmentally-coordinated approaches make fiscal sense.

9. Recommends increasing the collaborative spirit between DHE and NEASC with regards to the assessment of student learning.

10. Sees value for the Commonwealth and for individual campuses from the use of NSSE and CCSSE assessment tools and consortial arrangements within the framework of multiple assessment approaches and recommends further examination of the possibilities during Phase II activities of the Working Group on Student Learning Outcomes and Assessment.

Appendix B
LEAP: The Essential Learning Outcomes

Beginning in school, and continuing at successively higher levels across their college studies, students should prepare for twenty-first-century challenges by gaining:

Knowledge of Human Cultures and the Physical and Natural World

- Through study in the sciences and mathematics, social sciences, humanities, histories, languages, and the arts

Focused by engagement with big questions, both contemporary and enduring

Intellectual and Practical Skills, including

- Inquiry and analysis

- Critical and creative thinking

- Written and oral communication

- Quantitative literacy

- Information literacy

- Teamwork and problem solving

Practiced extensively, across the curriculum, in the context of progressively more challenging problems, projects, and standards for performance

Personal and Social Responsibility, including

- Civic knowledge and engagement—local and global

- Intercultural knowledge and competence

- Ethical reasoning and action

- Foundations and skills for lifelong learning

Anchored through active involvement with diverse communities and real-world challenges

Integrative Learning, including

- Synthesis and advanced accomplishment across general and specialized studies

Demonstrated through the application of knowledge, skills, and responsibilities to new settings and complex problems

Taken from Association of American Colleges and Universities, "College Learning for the New Global Century," p. 3.

Appendix C
LEAP Value Rubrics

Inquiry and Analysis

The ability to know when there is a need for information, to be able to identify, locate, evaluate, and effectively and responsibly use and share that information for the problem at hand.

Critical Thinking

Critical thinking is a habit of mind characterized by the comprehensive exploration of issues, ideas, artifacts, and events before accepting or formulating an opinion or conclusion.

Creative Thinking

Creative thinking is both the capacity to combine or synthesize existing ideas, images, or expertise in original ways and the experience of thinking, reacting, and working in an imaginative way characterized by a high degree of innovation, divergent thinking, and risk taking.

Written Communication

Written communication is the development and expression of ideas in writing. Written communication involves learning to work in many genres and styles. In can involve working with many different writing technologies, and mixing texts, data, and images. Written communication abilities develop through iterative experiences across the curriculum.

Oral Communication

Oral communication is a prepared, purposeful presentation designed to increase knowledge, to foster understanding, or to promote change in the listeners' attitudes, values, beliefs, or behaviors.

Reading

Reading is "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (Snow et al., 2002).

(From www.rand.org/pubsresearch_briefs/RB8024/index1.html)

Quantitative Literacy

Quantitative Literacy (QL)-also known as Numeracy or Quantitative Reasoning (QR)-is a "habit of mind," competency, and comfort in working with numerical data. Individuals with strong QL skills possess the ability to reason and solve quantitative problems from a wide array of authentic contexts and everyday life situations. They understand and can create sophisticated arguments supported by quantitative evidence and they can clearly communicate those arguments in a variety of formats (using words, tables, graphs, mathematical equations, etc., as appropriate).

Information Literacy

The ability to know when there is a need for information, to be able to identify, locate, evaluate, and effectively and responsibly use and share that information for the problem at hand.-The National Forum on Information Literacy

Teamwork

Teamwork is behaviors under the control of individual team members (effort they put into team tasks, their manner of interacting with others on team, and the quantity and quality of contributions they make to team discussions.)

Problem Solving

Problem solving is the process of designing, evaluating, and implementing a strategy to answer an open-ended question or achieve a desired goal.

Civic Engagement

Civic engagement is "working to make a difference in the civic life of our communities and developing the combination of knowledge, skills, values, and motivation to make that difference. It means promoting the quality of life in a community, through both political and non-political processes." (Excerpted from Civic Responsibility and Higher Education, edited by Thomas Ehrlich, published by Oryx Press, 2000, Preface, page vi). In addition, civic engagement encompasses actions wherein individuals participate in activities of personal and public concern that are both individually life enriching and socially beneficial to the community.

Intercultural Knowledge and Competence

Intercultural Knowledge and Competence is "a set of cognitive, affective, and behavioral skills and characteristics that support effective and appropriate interaction in a variety of cultural contexts." (Bennett, J.M. 2008. Transformative training: Designing programs for culture learning. In Contemporary leadership and intercultural competence: Understanding and utilizing cultural

diversity to build successful organizations, ed. M.A. Moodian, 95-110. Thousand Oaks, CA: Sage.)

Ethical Reasoning

Ethical Reasoning is reasoning about right and wrong human conduct. It requires students to be able to assess their own ethical values and the social context of problems, recognize ethical issues in a variety of settings, think about how different ethical perspectives might be applied to ethical dilemmas, and consider the ramifications of alternative actions. Students' ethical self-identity evolves as they practice ethical decision-making skills and learn how to describe and analyze positions on ethical issues.

Foundations and Skills for Lifelong Learning

Lifelong learning is "all purposeful learning activity, undertaken on an ongoing basis with the aim of improving knowledge, skills and competence". An endeavor of higher education is to prepare students to be this type of learner by developing specific dispositions and skills (described in this rubric) while in school. (From The European Commission. 2000. Commission staff working paper: A memorandum on lifelong learning. Retrieved September 3, 2003, from www.see-educoop.net/education_in/pdf/lifelong-oth-enl-t02.pdf.)

Integrative Learning

Integrative learning is an understanding and a disposition that a student builds across the curriculum and cocurriculum, from making simple connections among ideas and experiences to synthesizing and transferring learning to new, complex situations within and beyond the campus.

Source: Association of American Colleges and Universities. (2010). Assessing outcomes and improving achievement: Tips and tools for using rubrics. ed., T. L. Rhodes, Washington DC: Association of American Colleges and Universities.

Appendix D
Composite Indicator of Educational Capital

Over the course of our inquiry, the Working Group spent time considering the development of a Composite Indicator of Educational Capital because we thought that it might be useful during the period required for development and implementation of more focused learning outcomes measures and the composite learning indicator. We had in mind a composite indicator that would include learning measures along with educational attainment, success and workforce measures that the public often view as indicators that learning outcomes have been achieved. Ultimately, the Working Group chose to recommend a more focused indicator of student learning but it is included here as something that might prove interesting for the Vision Project as a whole. We will first discuss the background for the idea in the context of research undertaken by Margaret Miller and Peter Ewell to develop a model to allow for state comparisons of learning outcomes and then describe the composite educational capital indicator in greater detail.

In Chapter One, we noted the difficulty that the "Measuring Up" Project, the state-by-state report card for higher education begun in 2000, experienced trying to grade the states on student learning. The difficulty was caused by inadequate data on learning and resulted in all states receiving a grade of Incomplete in the learning category. This situation prompted the creation in 2001 of the National Forum on College-Level Learning directed by Margaret Miller. With support from the Pew Charitable Trusts, a five-state demonstration project was initiated to develop and test a model that states could use to evaluate learning in a way that would make interstate comparisons possible (Miller & Ewell, 2005; Miller, 2006; Ewell, 2008). The forum project involved work in Illinois, Kentucky, Oklahoma, Nevada and South Carolina. The focus at the state level is an important feature of this work. Participating states were asked to collect information that would answer two questions:

- What are the knowledge and skills of the population that states have available to them for developing or sustaining a competitive economy and vital civic life?

- How do all state colleges and universities (public, private, not-for-profit and for-profit) contribute to the development of those skills and knowledge?

Miller and Ewell referred to this work as a "comprehensive assessment of the educational capital of a state." Their model is multi-faceted and incorporates a selection of key indicators related to learning, skill attainment, workplace competencies, critical thinking and overall knowledge of college graduates. For the learning-related components, the project used national assessments of adult literacy and specially administered tests of general intellectual skills—the then relatively new Collegiate Learning Assessment (CLA). The resulting report, published by The National Center for Public Policy and Higher Education, presents this composite indicator as a grouping of associated outcome measures with benchmark comparisons to relevant outcomes for the nation as a whole. On the basis of this work, Miller and Ewell concluded that college-level learning could be measured in a way that allows for state-by-state comparisons. Although the study did not have extensive impact on subsequent state policy and practice, the concept of educational capital remains an important one for public higher education in the 50 states.

In our discussions, Working Group members could see potential in the concept of educational capital for informing the public about higher education in Massachusetts. Educational attainment and employment information is usually considered by members of the public and other stakeholders to be a good indicator that learning outcomes have been achieved in college. In fact many state accountability reports use attainment data as the primary indicator of student academic achievement. Combining indicators of education attainment, student learning, and post-graduation employment could produce an illuminating view of the educational capital produced by public higher education institutions to the Commonwealth of Massachusetts. It fits with the aspirations and timing of the Vision Project.

The Working Group recognized that the Miller/Ewell model is not directly applicable for our purposes because of its base in public and private institutions and in all citizens in a state, its reliance on a single test, and its use of data on adult literacy [National Assessment of Adult Literacy (NAAL's)], which is no longer kept current by the federal government. We wanted to pursue the idea, however, and in a November 2010 meeting with Peter Ewell we discussed the potential of developing a similar composite indicator for public higher education in Massachusetts. Although Ewell recognized that this would be neither a comprehensive assessment of the educational capital of Massachusetts, nor a composite student learning indicator, he thought the idea had merit. The Working Group continued to pursue it, thinking of it as an interim approach to statewide learning outcomes assessment for the Vision Project. We did not envision it as a replacement for a multi-measured dashboard presentation of student learning or a composite learning indicator, but thought it could be used as an additional feature for statewide reporting. Several limitations of the Composite Educational Capital Indicator should be noted first:

- It merges information from several different outcome areas in the Vision Project schema and thus complicates the effort to create mutually exclusive outcome categories and metrics.

- The concepts of educational capital and student learning are not synonymous. The underlying assumption is that they are both dependent on quality higher education and thus linked and may be combined to collectively indicate a measure of learning. It is not the best measure of learning, but it is one that could be developed from available data while allowing for additions over time as new measures become available.

- Composite indicators make it difficult to show change. The component measures and outcomes usually have some measures exhibiting downward or stagnant trends while others are trending upward. Consequently, it can be difficult to significantly "move the needle" and substantiate advancements that have been achieved through effective institutional policies and practices.

While the working group recognized the limitations, we continued to feel that the idea had merit for the Vision Project and that the following measures of learning and attainment could be included:

*Measures in Composite Indicator of Educational Capital*

- Performance on graduate school entrance exams.
- Performance on professional licensure exams.
- Transition from remedial course work into college-level courses.

- Successful completion of "gateway courses."
- Student Engagement Measures from NSSE and CCSSE surveys.
- Educational attainment and success.
- Post-graduation employment.
- Common measures from LEAP Embedded Assessments.

Benchmarking and Measuring Progress

The composite indicator presents certain challenges with regard to establishment of benchmarks and a suitable approach for measuring progress. But it gains strength from the fact that, while any of its component measures may not be meaningful or may be misleading if presented alone, each has greater assessment value if viewed in the context of other measures. Also, an underlying assumption of the composite indicator is that all of component measures are highly associated and are relatively less likely to display mixed and divergent outcomes. Ideally, the composite indicator would be presented as an integrated dashboard of component measures, compared to a baseline that represents a set of distinct benchmarks.

Example:
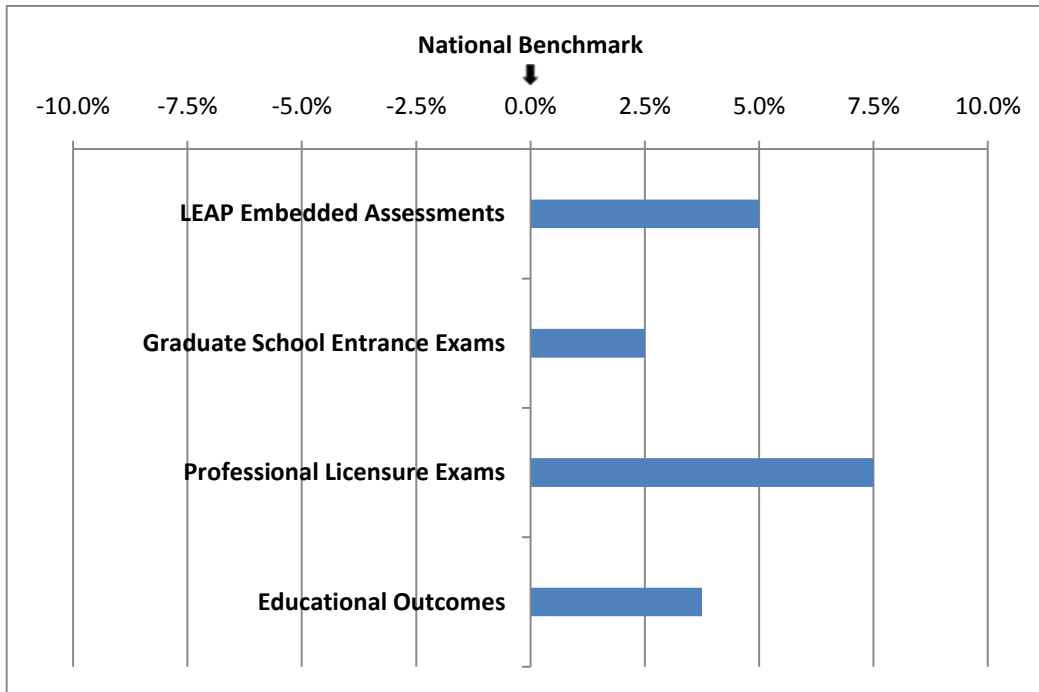
Composite Score

In addition to the integrated dashboard of component measures, a composite score may be presented which aggregates the individual educational capital outcomes and is compared to an analogous national composite score. This approach to the presentation of data enables us to summarize the assessment of learning across multiple indicators and to portray a picture of educational capital taken as a whole. There are a number of potential benefits and disadvantages to this approach. A poorly constructed composite score may result in the masking or obscuring of important differences and relations among component measures. This potential drawback may be mitigated by applying a sensible and conscientious method of selecting, weighting and benchmarking each component aspect of the composite score. A chief benefit of the composite score is that it essentially captures and addresses a broad spectrum of advantages that our higher education institutions provide to the Commonwealth.

References

Aldeman, C., & Carey, K. (2009). *Ready to assemble: Grading state higher education accountability systems.* Washington, DC: Education Sector.

Aldeman, C., & Kelly, A. P. (2010). *False fronts? Behind higher education's voluntary accountability systems.* Washington, DC: Education Sector.

Arum, R. & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses.* Chicago, IL: University of Chicago Press.

Arum, R., Roksa, J., & Cho, E. (2011). *Improving Undergraduate Learning: Findings and Recommendations from the Social Science Research Council – Collegiate Learning Assessment Longitudinal Project.* Washington, DC: Social Science Research Council.

Arum, R. (2011, February 3). *Prepared testimony.* Testimony before National Advisory Committee on Institutional Quality and Integrity, Alexandria, VA. Retrieved from http://highered.ssrc.org/wp-content/uploads/2011/02/Arum-Feb-3-NACIQI-comments.pdf.

Association of American Colleges and Universities. (2007). *College learning for the new global century.* Washington, DC: Association of American Colleges and Universities.

Association of American Colleges and Universities. (2008). *Our students' best work: A framework for accountability worthy of our mission.* (2nd ed.). Washington, DC: Association of American Colleges and Universities.

Association of American Colleges and Universities. (2011). *The LEAP vision for learning: Outcomes, practices, impact and employers views.* Washington, DC: Association of American Colleges and Universities.

Association of American Colleges and Universities & The Council for Higher Education Accreditation. (2008). *New leadership for student learning and accountability: A statement of principles, commitments to action.* Washington, DC: Association of American Colleges and Universities & the Council for Higher Education Accreditation.

Astin, A. W. (1991*). Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education*. New York: American Council on Education/Macmillan.

Banta, T. W. et al. (1996). *Assessment in practice: Putting principles to work on college campuses.* San Francisco: Jossey-Bass.

Banta, T. W., Griffin, M., Flateby, T., & Kahn, S. (2009). *Three promising alternatives for assessing college students' knowledge and skills.* (NILOA Occasional Paper No. 2) Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Blaich, C. F. (2007). *Overview of findings from the first year of the Wabash National Study of liberal arts education.* Wabash College, Center of Inquiry in the Liberal Arts. Retrieved from http://www.liberalarts.wabash.edu/research/.

Blaich, C. F., & Wise, K.F. (2011). *From gathering to using assessment results: Lessons from the Wabash National Study.* (NILOA Occasional Paper No. 8) Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Brownell, J. E., & Swaner, L. E. (2010). *Five high-impact practices: Research on learning outcomes, completion, and quality.* Washington, DC: Association of American Colleges and Universities.

Callan, P. (2006). *Commentary on Measuring Up 2006 report.* San Jose, CA: The National Center for Public Policy and Higher Education.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006) Student engagement and student learning: Testing the linkages." *Research in Higher Education 47, 1- 32.*

Eaton, J. S. (2011). Federalizing accreditation: A quandary for higher education. *Inside Accreditation, 7 (1), Feb 16.* Retrieved from http://www.chea.org/ia/images/IA_banner2.jpg. Washington, DC: Council for Higher Education Accreditation.

Esposito, G., Mantil, A., Pearson, R., & St. Julien, T. (2010). *Assessment of student learning outcomes in higher education: Practices and lessons from 14 states.* Unpublished manuscript. Paper prepared for Working Group by students in Professor Reville's class (A-125: State Education Policy) at the Harvard Graduate School of Education, Cambridge, MA.

Ewell, P. T. (2008). Assessment and accountability in America today: Background and context. In V.M.H. Bordon and G. Pike (Eds), *Assessing and accounting for student learning: Beyond the Spellings Commission.* (New Directions for Institutional Research, Assessment Supplement 2007, p. 7-18). San Francisco, CA. Jossey-Bass.

Ewell, P. T. (2009). *Assessment, accountability and improvement: Revisiting the tension.* (NILOA Occasional Paper No. 1) Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Ewell, P. T., & Jones, D. P. (1985). *Costs of assessment.* Boulder, CO: National Center for Higher Education Management Systems.

Ewell, P., Jankowski, N., & Provesis, S. (2010). *Connecting state policies on assessment with institutional assessment activity.* NILOA State Study. Retrieved from www.learningoutcomeassessment.org/documents/NILOAStateStudy_000pdf

Finley, A. "Draft concept paper for Compass II." (S. Albertine, AAC&U, personal communication, August, 2010.

Gawande, A. (2007). *Better: A surgeon's notes on performance.* New York: Metropolitan Books, Henry Holt and Company.

Gawande, A. (2011, January 24). The hot spotters. *The New Yorker*.

Gladwell, M. (2011, February 14&21). The order of things: What college rankings really tell us. *The New Yorker*, 68-75.

Harper, V.B. Jr. (2009). Virginia's value added: A diverse system perspective. *Assessment Update.* 21, 4.

Hart Research Associated, Inc. *How should colleges assess and improve student learning? Employers views on the accountability challenge.* Washington, DC: Peter D. Hart Research Associated, Inc., 2008.

Klein, S., Liu, O. L., & Sconing, J. (2009). *Test validity study (TVS) report.* Retrieved from http://www.voluntarysystem.org/docs/reports/TVSReport_Final.pdf.

Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter.* Washington, DC: Association of American Colleges and Universities.

Kuh, G., & Ikenberry, S. (2009). *More than you think, less than we need: learning outcomes assessment in American higher education.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA). http://www.learningoutcomeassessment.org/NILOAsurveyresults09.htm.

McClenney, K., Marti, C. N., & Adkins, C. (2006). *Student engagement and student outcomes: Key findings from CCSSE validation research.* http://www.ccsse.org/about

McPherson, P., & Shulenburger, D. (2006-August). *Toward a public universities and colleges Voluntary System of Accountability for undergraduate education (VSA): A NASULGC and AASCU discussion draft.* Washington, DC. National Association of State Universities and Land Grant Colleges (NASULGC).

Miller, M. A. (2006, May) *Policy alert: Assessing college-level learning.* San Jose, CA: The National Center for Public Policy and Higher Education.

Miller, M. A., & Ewell, P. T. (2005). *Measuring up on college-level learning.* San Jose, CA: The National Center for Public Policy and Higher Education.

National Center for Public Policy in Higher Education. (2000-2010). *Measuring up: The fifty-state report card on higher education.* San Jose, CA: National Center for Public Policy and Higher Education.

National Commission on Accountability in Higher Education. (2005). *Accountability for better results: A national imperative for higher education.* Denver, CO: State Higher Education Executive Officers.

National Institute for Learning Outcomes Assessment. (2011). *Providing evidence of student learning: A Transparency Framework. Retrieved from www.learningoutcomeassessment.org/TFexamples.htlm.*

National Survey of Student Engagement (NSSE). (2007). *Experiences that matter: Enhancing student learning and success.* Bloomington, IN: Center for Postsecondary Research, Indiana University Bloomington.

New England Association of Schools and Colleges (NEASC), Commission on Institutions of Higher Education (2005). *Standards for accreditation.* Retrieved from http://cihe.neasc.org/downloads/Standards/Standards_for_Accreditation_2006.pdf

New England Association of Schools and Colleges (NEASC), Commission on Institutions of Higher Education. (2008, August). *Student achievement and success.* Bedford, MA: Author. Retrieved from http://cihe.neasc.org/downloads/studentachievement.

Pascarella, E. T., Seifert, T. A., & Blaich, C. F. (2007, January/February). How effective are the NSSE benchmarks in predicting important educational outcomes?" <u>*Change*</u>.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research.* San Francisco, CA: Jossey-Bass.

Rhodes, T. L. (Ed.). (2010). *Assessing outcomes and improving achievement: Tips and tools for using rubrics.* Washington, DC: Association of American Colleges and Universities.

Shavelson, R. J., (2010). *Measuring college learning responsibly: Accountability in a new era.* Palo Alto: Stanford University Press.

Shulenburger, D. and Keller, C., (2009). *Interpretation of findings of the test validity study conducted for the Voluntary System of Accountability*. Washington, DC: Voluntary System of Accountability. Available from http://www.voluntarysystem.org.

Shulman, L.S. (2007). "Counting and recounting: Assessment and the quest for accountability." *Change,* 39 (1), 20-25*.*

Swing, R. L., & Coogan, C. S. (2010, May). *Valuing assessment: Cost-benefit considerations. .* (NILOA Occasional Paper No. 5) Urbana, IL: University for Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

U. S. Department of Education. (2006) *A test of leadership: Charting the future of U.S. higher education.* (Report of the commission appointed by Secretary of Education Margaret Spellings). Washington, DC: U.S. Department of Education.

Walvoord, B. E., (2010). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. (2nd ed.). San Francisco: Jossey-Bass.

Wellman, J.V. (2010, January). *Connecting the dots between learning and resources.* (NILOA Occasional Paper No. 3) Urbana, IL: University for Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Materials referenced in Phase One Report but Not Used for Phase Two (Materials used in both phases appear above)

Association of American Colleges and Universities and Association of Public Land Grant Universities. (2010). *Rising to the challenge: Meaningful assessment of student learning.* Washington, DC: Association of American Colleges and Universities.

Banta, T.W. (2007). Can assessment for accountability complement assessment for improvement? *peer Review*, Spring, 9-12.

Banta, T. W. (2008). Trying to clothe the emperor. *Assessment Update*, 20, 2 (March-April 2008) 3-4, 15-16. Wiley Periodicals.

Benjamin, R. (2009). *Returning to learning in an age of assessment.* New York: Council for the Advancement of Education.

Bok, D. (2006). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more.* Princeton, NJ: Princeton University Press, 2006.

Council for the Advancement of Education. (2009). *CAE Board statement on appropriate uses of the CLA.* New York: Council for the Advancement of Education. Retrieved from http://www.collegiatelearningassessment.org.

Dwyer, Millett and Payne. (2006). *A culture of evidence: postsecondary assessment and learning outcomes.* Culture of Evidence Series. Princeton, NJ: Educational Testing Service.

Massachusetts Board of Higher Education. (2009). *Final report of the Commissioner's Advisory Group on Undergraduate Education.* N. Hoffman, chair. Available at: http://www.mass.edu.

Miller, R. (2007). *Assessment in cycles of improvement: faculty designs for essential learning outcomes.* Washington, DC: Association of American Colleges and Universities.

Millett, Stickler, Dwyer and Payne. (2007). *A culture of evidence: critical features of assessments for postsecondary student learning.* (Culture of Evidence Series). Princeton, NJ: Educational Testing Service.

Millett, Payne, Dwyer, Stickler and Alexiou. (2008). *A culture of evidence: an evidence-centered approach to accountability for student learning outcomes.* (Culture of Evidence Series). Princeton, New Jersey: Educational Testing Service.

Rosnik, P. "Beware the Slippery Slope of Outcomes Assessment," and "Mathematics In Search of Meaning" (Personal communication, 2010).